RWLP0078

INTERNATIONAL AGENCY FOR RESEARCH ON CANCER
WORLD HEALTH ORGANIZATION

ch on Cancer

C) was established in 1965 by the
anization within the framework of
Agency are at Lyon, France.

rating particularly on the epidemi-
the human environment. Its field
rch carried out in the Agency's lab-
ements, in national research insti-
rogramme for the education and

te to the dissemination of author-
Information about IARC publica-
t: http://www.iarc.fr/

Quantitative Estimation and Prediction of Human Cancer Risks

Edited by
S. Moolgavkar, D. Krewski, L. Zeise,
E. Cardis and H. Møller

IARC Scientific Publications No. 131

International Agency for Research on Cancer
Lyon, France
1999

PM3006483527

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Published by the International Agency for Research on Cancer,
150 cours Albert Thomas, 69372 Lyon cédex 08, France

© International Agency for Research on Cancer, 1999

Distributed by Oxford University Press, Walton Street, Oxford OX2 6DP, UK
(fax: +44 1865 267782) and in the USA by Oxford University Press, 2001 Evans Road, Carey,
NC 27513 (fax: +1 919 677 1303). All IARC publications can also be ordered directly from IARC*Press*
(fax: +33 04 72 73 83 02; E-mail: press@iarc.fr).

Publications of the World Health Organization enjoy copyright protection in accordance with the provisions of Protocol 2 of the Universal Copyright Convention. All rights reserved.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the World Health Organization concerning the legal status of any country, territory, city, or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

The authors alone are responsible for the views expressed in this publication.

The International Agency for Research on Cancer welcomes requests for permission to reproduce or translate its publications, in part or in full. Applications and enquiries should be addressed to the Editorial & Publications Service, International Agency for Research on Cancer, which will be glad to provide the latest information on any changes made to the text, plans for new editions, and reprints and translations already available.

IARC Library Cataloguing in Publication Data

Quantitative estimation and prediction of human cancer risks /
editors, S. Moolgavkar ... [et al.]

(IARC scientific publications ; 131)

1. Neoplasms – epidemiology
2. Risk Assessment I. Moolgavkar, Suresh H
II. Series

ISBN 92 832 2131 1 (NLM Classification W1)
ISSN 0300-5085

Printed in France

Foreword ....

Contributors ....

Acknowledgements ....

Chapter 1. Quantitative estin and role in cancer prev

A.J. McMichael and A. We

1.1 Introduction ....
1.2 Historical background ....
1.3 Definitions of terms and ....
1.4 Regulation and control ....
1.5 Risk prediction ....
1.6 Conclusions ....
References ....

Chapter 2. Quantitative esti review of existing acti

L. Zeise, E. Cardis, K. Her

2.1 Introduction ....
2.2 QEP by international bo
2.2.1 World Health Organiza
2.2.1.1 Drinking-water ....
2.2.1.2 Ambient air ....
2.2.1.3 Food residues and
2.2.1.4 General environme
2.2.1.5 Developments in th
2.2.2 International bodies th
2.2.3 Other international bo
2.3 QEP by national institut
2.3.1 Netherlands ....
2.3.2 USA ....
2.3.2.1 Default risk predict
2.3.2.2 Use of QEPs ....
2.3.3 Canada ....
2.3.4 Norway ....
2.3.5 United Kingdom ....
2.3.6 Germany ....
2.3.7 Denmark ....
2.4 Standard approaches to
2.5 Attempts to improve Q
2.5.1 "Biologically based" a
2.5.1.1 Integrated modelli dynamics: 2,3,7,8

CMH

PM3006483528

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Quantitative Estimation and Prediction of Human Cancer Risks
S. Moolgavkar, D. Krewski, L. Zeise, E. Cardis and H. Moller
IARC Scientific Publications No. 131
International Agency for Research on Cancer, Lyon, 1999

4: Measurement of Exposure and Outcome in Epidemiological Studies used for Quantitative Estimation and Prediction of Risk

Bruce Armstrong and Paolo Boffetta

4.1 Introduction

Making the right measurements of the exposure (including agents confounded with it and modifiers of its effect) and of the outcome of interest (usually cancer, a preneoplastic lesion or another intermediate step between exposure and cancer), and making them accurately, are crucial to the valid quantitative estimation and prediction (QEP) of cancer risk. These requirements are the two sides of the same coin: if the right measurements are not made, the consequences for QEP will be the same as if inaccurate measurements are made: the dose–response relationship will be biased and predictions made from it may be far from the truth. They are both errors in measurement.

Theoretically, error in measurement of exposure and outcome will have similar effects on the dose–response relationship. In the simplest kind of study, investigating the association between a binary exposure variable and a binary outcome variable in the absence of confounding or other intervening variables, error in measurement of either the exposure or the outcome variable has the same effect on the estimate of association between exposure and outcome. In practice, however, the measurement of exposure is usually more complex than that of outcome, mainly because of the problem of measuring past exposures, the need to measure exposure to several agents (the agent of interest, confounders and effect modifiers), and the multiplicity of methods available to measure exposure. As a result, greater attention is usually paid to the correct measurement of exposure than to that of outcome. In addition, the study of the effect of exposure measurement error has presented more interesting theoretical problems than that of outcome measurement error, and has received more attention in the relevant literature.

In principle, error in the measurement of exposure is a concern only for observational research. Error in measurement of outcome, on the other hand, is probably of equal concern in both observational and experimental research. In well designed and conducted experiments, the exact exposure of each subject to the agent should be known and controlled by the investigator, and randomization should prevent confounding. In practice, it is usually desirable to measure key confounding variables and control for them in the analysis because of the possibility that, by chance, important confounding may be present. In addition, measurement error again becomes an issue when, in an experimental study, it is desirable to estimate exposure within the body at the level of the target tissue or target molecules. In these experimental situations, the problems and issues related to error in measurement of exposure are the same as they are in observational research.

In this chapter, we shall consider in detail aspects of the measurement of exposure from the perspective of observational epidemiological research; we shall outline the exposure measurements that should be made, the effects of error in exposure measurement on QEP, the prevention of error in exposure measurement and the control of its effects when it has not been prevented. These issues will also be considered in a more limited way for outcome measurement in epidemiological studies. The book *Principles of Exposure Measurement in Epidemiology* (Armstrong *et al.*, 1992) is acknowledged as a major source of material for this chapter.

PM3006483529

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

4.2 Making the right exposure measurements

4.2.1 Nature of exposure measurements

Exposure variables should be *specific* in the sense that they measure indivisible agents of exposure. Thus, for example, it is better to measure the different forms of tobacco smoking (cigarettes, pipes, cigars, etc.) separately than generically as "smoking". The dose–response relationship of smoking-related lung cancer differs by method of smoking (see, for example, Higgins *et al.*, 1988) and use of a non-specific variable will lead to a dose–response relationship that cannot readily be interpreted in terms of any individual exposure and therefore used to undertake specific QEP. Similarly, the exposure variables measured should permit the distinction of the exposure of interest from possibly confounding exposures. In this regard, it is relevant to note that, in evaluations of the evidence for the carcinogenicity of chemicals in humans made by the International Agency for Research on Cancer, the commonest reason for applying a "limited" or "inadequate" evidence classification was that exposure to the agent of interest had not been distinguished from exposure to other potentially carcinogenic agents in the environment in which it had occurred (Armstrong, 1985). If such data cannot support qualitative assessment of risk, they will be even less useful for QEP.

The exposure variables should also be *sensitive* in the sense that they include all ways by which subjects may be exposed to the agent of interest. For example, in examining the relationship between environmental tobacco smoke (passive smoking) and lung cancer a number of studies in non-smoking women have used the spouses smoking habits as the sole exposure variable. This variable excludes home exposures prior to marriage, home exposures other than to the spouse's smoke, and workplace, social and other exposures outside the home at any time. The importance of multiple sources of exposure to environmental tobacco smoke has been shown in recall of lifetime exposure (Cummings *et al.*, 1989) and in comprehensive analyses of the determinants of current exposure (Coghlin *et al.*, 1989; Riboli *et al.*, 1990). A complete history of exposure is necessary, therefore, to obtain accurate measurements of exposure to environmental tobacco smoke. Variation in the error caused by incomplete (or insensitive) measurement of passive smoking may explain some of the variation in the strength of the relationship that has been observed between it and lung cancer in non-smoking women (see, for example, Garfinkel, 1981; Hirayama, 1981). More recent studies have attempted more comprehensive coverage of passive exposure to tobacco smoke (see, for example, Fontham *et al.*, 1991; Stockwell *et al.*, 1992).

4.2.2 Measurement of dose

For QEP, the dose should be measured in quantitative terms, preferably as a dose rate (rather than the dose accumulated over, e.g. the whole of the period of exposure), in the most fundamental units in which the agent is usually measured. For example, to obtain an accurate quantitative measure of dose of the combined oral contraceptive pill, a complete history covering the nature and period of use of each kind of pill used should be obtained. Provided the exact formulations are identified (and this is not impossible; UK National Case–Control Study Group, 1989), not only can use of combined oral contraceptives be measured in terms of, say, years of use but also in actual weights of the oestrogenic and progestagenic components consumed. Information on dose rate would allow measurement of intake in specific periods of time within the whole exposure period (see below).

A number of levels at which dose may be measured are important to QEP (Figure 4.1). The *exposure,* sometimes called the available dose, is measured in the environment external to the subject, e.g. the concentration of asbestos fibres per ml of ambient air over a small time interval. The exposure is often the measurement used for regulatory purposes, and may therefore be the quantity most relevant to QEP. It does not usually coincide with the *administered dose* or intake, i.e. the actual amount of the agent coming into contact with the human body. How much of the exposure becomes administered dose depends on the subject's physiology and behaviour, e.g. the respiratory volume at rest and during activity and the quantities of food, drinks, and medications actually ingested. From a biological viewpoint, even the administered dose can usually only be regarded as a proxy measurement of the *absorbed dose* or uptake, i.e. the dose that actually enters the body.

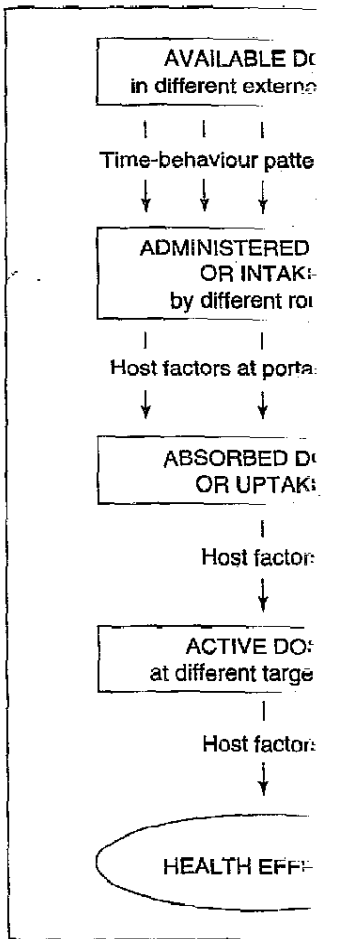


Figure 4.1. Levels at which dose may

The absorbed dose, in turn, is
that really matters in terms of
the *active or biologically effecti*
the body (organs, tissues, cell
are the specific targets of acti

It could be argued that, fc
tive estimation of risk, the m
to measure would be the b
dose at the level of interacti
targets (e.g. specific DNA a
specific mutations) because it

PM3006483530

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

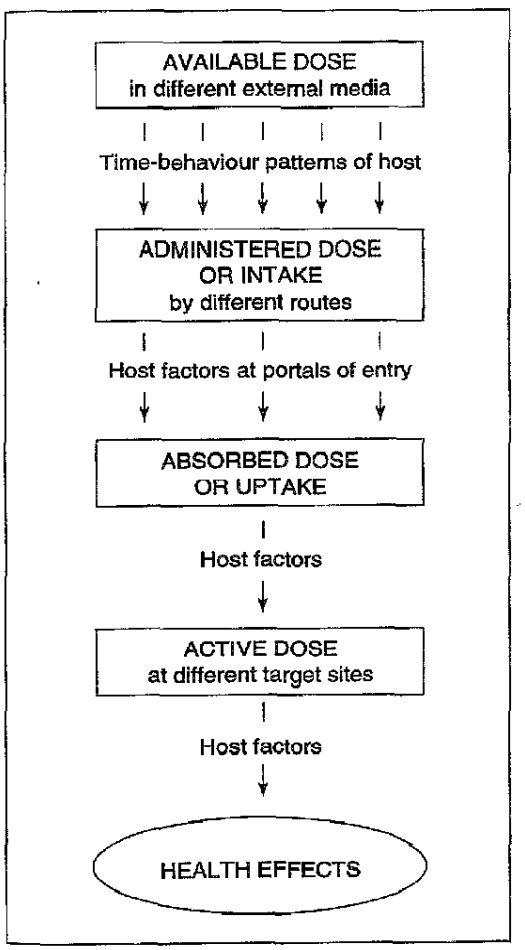


Figure 4.1. Levels at which dose may be measured

The absorbed dose, in turn, is a proxy for the dose that really matters in terms of causing the disease, the *active or biologically effective dose* at the sites in the body (organs, tissues, cells, molecules) which are the specific targets of action of the agent.

It could be argued that, for accurate quantitative estimation of risk, the most appropriate dose to measure would be the biologically effective dose at the level of interaction with the cellular targets (e.g. specific DNA adducts or exposure-specific mutations) because it is measured after all preceding sources of variation (environmental conditions influencing available dose, behaviours affecting the administered dose, physiological factors affecting the absorbed dose and metabolic factors affecting the active dose) have been taken into account. On the other hand, available dose is the measurement most commonly (although not invariably) subject to public health regulation and therefore the one for which QEP is required. There is, then, a conflict between what might be best for estimation and what is needed for prediction.

If, as may be hoped, measurements of biologically effective dose are able to fulfill their promise as accurate, quantitative predictors of risk of human cancer, at least in some circumstances, it may be that measurement of the biologically effective dose of some agents in exposed humans may become the most appropriate variable for public health regulation. In the meantime, it will still be important to strive for accurate measurement of available dose as well as to develop and evaluate the predictive capacities of absorbed and biologically effective doses and to understand better the relationships between these various levels of dose.

4.2.3 Measurement of variation in exposure with time

As far as possible, each exposure should be characterized by when it first began, when it finally ended (if at all), and how the dose rate varied during the period of exposure. The time relationships of exposure are important for several reasons. Thus, duration of exposure is a critical determinant of the total amount of exposure that has occurred. Then, for a cancer occurring at a particular time, there exists a "critical" (or effective) time window (the aetiologically relevant exposure period) during which exposure can be relevant to its causation (Rothman, 1981). Inclusion of exposure outside that time window will lead to error in the exposure measurement (a case of non-specificity in the terms outlined above), and therefore to error in the estimation of dose–response. We do not generally know when the effective time window for exposure occurred for any particular exposure–cancer combination. The matter is further complicated by the possibility of multiple effective time windows for an

PM3006483531

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

agent when, for example, it has both early- and late-stage effects. For epithelial cancers, it has been commonly assumed that the effective time window was at least 10 years ago, and exposure in the last 10 or so years is therefore excluded from the analysis. This approach ignores late-stage effects. An alternative approach, such as that exemplified by the "serially additive expected dose model" (Smith *et al.*, 1980), may be used to seek the location of the effective time window empirically. This allows for the possibility that more than one such time window will be found. Full data on the time relationships of exposure may allow the construction of more complete and potentially more informative mechanistic exposure–response models for QEP than are currently available (see Chapter 6).

Finally, pattern of exposure during the exposure period may also be important. For example, to explain the apparently anomalous occurrence of a higher incidence of malignant melanoma of the skin in indoor workers than in outdoor workers, it was postulated that a particular total dose of sunlight may be more effective in causing melanoma if it is received intermittently or irregularly rather than frequently or continuously (Holman *et al.*, 1983). There is empirical evidence supporting this hypothesis (Armstrong, 1988). Unless such pattern effects are taken into consideration, QEP may produce misleading results.

Measurement of duration, dose rate, and pattern of exposure will permit the maximum flexibility in defining exposure in a biologically relevant way when analysing the relationship between exposure and disease. Alternatively, if there is no prior hypothesis regarding which representation of dose and time would be the most appropriate, the representation that best predicts disease incidence may then be sought empirically. In practice, however, it may be difficult to distinguish between different, possibly appropriate representations of dose, especially when many different characteristics of dose are included in statistical dose–response models.

Lee-Feldstein (1989) has given an example of an empirical search for the most appropriate representation of dose in relation to time. She examined the quantitative relationships between exposure to arsenic in a copper smelter and risk of lung cancer in two cohorts of men, one first employed before 1925 and the other employed between 1925 and 1947. Exposure was defined in terms of job with maximum exposure to arsenic (a "peak" exposure definition), cumulative exposure based on arithmetic or geometric mean levels of arsenic in specific job areas, and time-weighted average exposures based on arithmetic or geometric means. Analyses were carried out with and without exclusion of exposure in the last 10 years before death. Table 4.1 shows the results obtained for both cohorts with three different definitions of expo-

Table 4.1 Relationship of risk of lung cancer with arsenic exposure under different definitions of exposure in two cohorts of copper smelters[a]

Definition of exposure	First cohort			Second cohort		
	$\hat{\beta}^b$	SE of $\hat{\beta}^b$	Likelihood ratio statistic for $\beta = 0$	$\hat{\beta}^b$	SE of $\hat{\beta}^b$	Likelihood ratio statistic for $\beta = 0$
Maximum exposure category	0.55	0.13	18.5	0.30	0.14	4.5
Cumulative exposure	0.68×10^{-4}	0.20×10^{-4}	10.4	0.65×10^{-4}	0.34×10^{-4}	3.2
Time-weighted average exposure	0.026	0.0075	10.8	0.021	0.0075	7.0

[a] Based on data from Lee-Feldstein (1989).
[b] $\hat{\beta}$ = estimated logistic regression coefficients for models in which each variable was fitted as a continuous variable. SE = standard error.

sure. The cumulative and exposures shown are those means since they gave sli ratio statistics than thos means. In the first coho ratio statistic in each case exposure statistically show tion with lung cancer, whe age and cumulative exposu association. In the secon time-weighted average app ly associated with lung c exposure, and the maxi was intermediate between this case, that some meas sure over time was import The values in Table 4.1 removing the last 10 yea exposure estimates.

With respect to time, it some measurement insu comparatively short perio therefore of limited use fo multiple measurements o over the aetiologically rele exposure can be considere stant over time. For exam concentration of asbestos i ly speaking, only to the pe the air sample was taken, cotinine reflects exposure within the past 3–4 days (measurement of adducts serum albumin reflects int the preceding several week *et al.*, 1990). Unless reco ments may be of little valu sure in a case–control stud aetiologically relevant exp been 20 or 30 years ago. Si ited value in prospective co are repeated at intervals du od. If repeated, they could ble data on dose rate an obtainable. Such data, ho tional problems, such as th the multiple measures in possibility that the errors vary with time, e.g. becaus nology used.

PM3006483532

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

tion, dose rate, and pat-
rmit the maximum flexi-
re in a biologically rele-
the relationship between
ternatively, if there is no
which representation of
he most appropriate, the
redicts disease incidence
rically. In practice, how-
distinguish between dif-
iate representations of
ny different characteris-
cluded in statistical

s given an example of an
st appropriate represen-
to time. She examined
ips between exposure to
and risk of lung cancer
e first employed before
yed between 1925 and
ed in terms of job with
nic (a "peak" exposure
posure based on arith-
evels of arsenic in spe-
eighted average expo-
or geometric means.
vith and without exclu-
10 years before death.
lts obtained for both
t definitions of expo-

der different

of $\hat{\beta}^b$	Likelihood ratio statistic for $\beta = 0$
	4.5
$\times 10^{-4}$	3.2
75	7.0

b. SE = standard error.

sure. The cumulative and time-weighted average exposures shown are those based on the arithmetic means since they gave slightly higher likelihood ratio statistics than those based on geometric means. In the first cohort, using the likelihood ratio statistic in each case as a criterion, maximum exposure statistically showed the strongest association with lung cancer, whereas time-weighted average and cumulative exposure were similar in their association. In the second cohort, however, the time-weighted average appeared to be more strongly associated with lung cancer than cumulative exposure, and the maximum exposure category was intermediate between the two. It appears, in this case, that some measure of intensity of exposure over time was important in determining risk. The values in Table 4.1 were little affected by removing the last 10 years of exposure from the exposure estimates.

With respect to time, it is relevant to note that some measurement instruments relate only to comparatively short periods of exposure and are therefore of limited use for QEP of cancer, unless multiple measurements can or have been taken over the aetiologically relevant exposure period or exposure can be considered to be reasonably constant over time. For example, measurement of the concentration of asbestos fibres in air relates, strictly speaking, only to the period of time over which the air sample was taken, measurement of urinary cotinine reflects exposure to tobacco smoke only within the past 3–4 days (Riboli *et al.*, 1990), and measurement of adducts of aflatoxin B_1 with serum albumin reflects intake of aflatoxin B_1 over the preceding several weeks to a few months (Wild *et al.*, 1990). Unless records exist, such measurements may be of little value in documenting exposure in a case–control study of cancer in which the aetiologically relevant exposure period may have been 20 or 30 years ago. Similarly, they are of limited value in prospective cohort studies unless they are repeated at intervals during the follow-up period. If repeated, they could provide the best possible data on dose rate and pattern of exposure obtainable. Such data, however, may bring additional problems, such as the non-independence of the multiple measures in any individual and the possibility that the errors in measurement may vary with time, e.g. because of changes in the technology used.

Methods that purport to be able to measure exposure over long periods of time also have limitations. First, even if accurate (as, for example, measurements of UV-specific mutations of the *p53* gene in normal skin might be for UV exposure; Nakazawa *et al.*, 1994), they may provide information on cumulative exposure only and tell us nothing about dose rate and pattern of exposure. Second, when based on human recall, as commonly they must be, they are usually of uncertain, if not doubtful, accuracy.

4.3 Effects of exposure measurement error on QEP

Measurement error can be simply conceived as follows:

$$X_i = T_i + b + E_i$$

where the observed measure, X_i, differs from the true value T_i by the systematic error or bias, b, which occurs, on average, in the measurements of all measured subjects, and the non-systematic error, E_i, that varies unpredictably from subject to subject (Armstrong *et al.*, 1992). X, T, and E are variables with distributions; their expectations are denoted by μ_X, μ_T and μ_E and their variances by σ^2_X, σ^2_T and σ^2_E respectively. Because the average measurement error in X is expressed as a constant, b, it follows that μ_E, the population mean of the non-systematic measurement error, is assumed to be zero.

When two or more groups of subjects are being compared, measurement error can be either differential or non-differential. It is differential if the value of b (the bias) differs between the groups or if the precision of X (the observed measurements) differs between the groups (due to differences between the groups in the value of σ^2_E), or if both are true. If neither is true, the error is non-differential.

4.3.1 Effect of exposure measurement error on the dose–response relationship

The effect of error in the measurement of the primary exposure on the empirical dose–response relationship depends on (i) the type of study, namely whether it is analytical (cohort or case–control) or ecological; (ii) the type of measurement, namely whether it is continuous or categorical; and (iii) whether the error is differential or non-differential.

PM3006483533

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

4.3.1.1 Analytical epidemiological studies

In general terms, it can be said that, when a continuously distributed measure of exposure is related to disease outcome in a logistic model and there is differential error in measurement of the exposure, the observable odds ratio (OR_O) for the disease, per unit of exposure, bears no predictable relationship to the true odds ratio (OR_T). In other words, it may be closer to the null value, further from the null value, or on the other side of the null value (e.g. below 1.0 instead of above 1.0) from the true value.

If error in measurement of a continuous variable is non-differential its effects may be more predictable. If it is assumed that the degree of measurement error is uncorrelated with the true value of the measurement and the true values are normally distributed and have the same variance in subjects with and without disease, it can be shown that non-differential error in exposure measurement causes flattening of the curve of the relationship of the odds ratio with exposure towards the null value (odds ratio constant at 1.0 with increasing exposure; see Figure 4.2). The relationship between the observable logistic regression coefficient, β_O (which is related to the observable odds ratio for u units of exposure as follows, $OR_O = \exp(\beta_O u)$, is:

$$\beta_O = \rho^2_{TX}\beta_T$$

In this expression, T and X are the true and observed values of the measurement, respectively, and ρ is the coefficient of correlation of T with X, the validity coefficient, and β_T is the true logistic regression coefficient. Thus β_O, and therefore OR_O, falls as the value of ρ falls. For example, when ρ has a value of 0.5, an OR_T of 2.0 for one unit of exposure is attenuated substantially to an OR_O of 1.19.

When the assumptions of the preceding paragraph are not true, the odds ratio curve shown in Figure 4.2 can take a variety of shapes including convex upwards, convex downwards with the lowest odd ratios falling below the null value, or even sigmoid.

When the exposure variable is categorical, the effects of measurement error on the exposure–disease and dose–response relationships

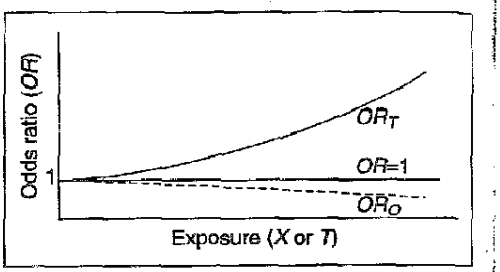


Figure 4.2. Relationship between observable logistic regression coefficient and true logistic regression coefficient for different values of true and observed measurement (see text for details)

are analogous to those when the variable is continuous. If there is differential error in measurement, the observable odds ratio (or odds ratios if there are more than two exposure categories) may bear almost any relationship to its true value.

If there is non-differential measurement error, the effect depends on whether there are two or more than two exposure categories. When there are two exposure categories, OR_O will be biased towards the null value relative to OR_T but will not cross over the null value provided that the measurement classifies a truly exposed person as exposed with the same or a greater probability than it classifies a truly unexposed person as exposed.

The position is more complex and less predictable when there are more than two exposure categories (Birkett, 1992). Briefly, when risk of disease truly increases monotonically with increasing exposure, the OR_O for the most extreme exposure category will always be biased towards the null while that for intermediate categories may be biased away from it. Values of OR_O for lower levels of exposure can cross over the null value, and all values of OR_O can be so biased that the apparent direction of the relationship can be reversed. If it is assumed that misclassification occurs only between adjacent exposure categories, the overall direction of the relationship cannot be reversed. However, OR_O for the lowest category can cross over the null, and a monotonic dose–response relationship may thus appear as a U- or J-shaped curve.

4.3.1.2 Ecological studies

The effect of meas
dose–response estimation in
only recently been considere
For an ecological study in wh
sured by the proportion of in
ulation who were exposed, it
estimated relative risks was ak
than towards the null value a
be quite large. It appears, h
exposure of each population i
common measure (e.g. area a
be the most useful for the qua
risk, the bias would usually
value (Brenner *et al.*, 1992). T
not been rigorously worked o

4.3.2 Measurement error in c

Error in the measurement
able influences the re
control of the effects of
dose–response relationship of
variable with disease.

4.3.2.1 Analytical epidemiol

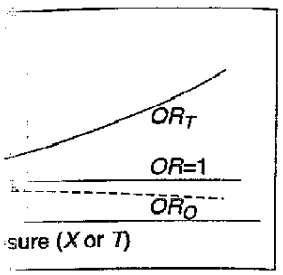
When the confounder is m
ferential error and the exposu
ly, the effects of the confound
ly controlled and the effects
pendent of the confounder,
smaller than they really are, de
tion of confounding. When
sured with non-differential
founder is measured perfec
ease–exposure association ma
the null value to an even gr
crude association. When boi
error, the adjusted disease–
may, with reference to the
association, be biased toward
null.

4.3.2.2 Ecological studies

Confounding in ecological
plex than in analytical studies,
tially much greater, and its c
(Greenland & Morgenstern,
1992). The effects of error in
confounders on the results of
not been described

PM3006483534

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001



ween observable logistic regres-
tic regression coefficient for dif-
erved measurement (see text for

se when the variable is
is differential error in
servable odds ratio (or
more than two exposure
most any relationship to

fferential measurement
ds on whether there are
vo exposure categories.
xposure categories, OR_O
he null value relative to
over the null value pro-
ement classifies a truly
sed with the same or a
n it classifies a truly
posed.
complex and less pre-
e more than two expo-
1992). Briefly, when
ncreases monotonically
e, the OR_O for the most
egory will always be
while that for interme-
biased away from it.
levels of exposure can
, and all values of OR_O
e apparent direction of
e reversed. If it is
ification occurs only
re categories, the over-
lationship cannot be
for the lowest category
ll, and a monotonic
ip may thus appear as

4.3.1.2 Ecological studies

The effect of measurement error on dose–response estimation in ecological studies has only recently been considered (Brenner *et al.*, 1992). For an ecological study in which exposure was measured by the proportion of individuals in each population who were exposed, it was shown that bias in estimated relative risks was always away from, rather than towards the null value and that the bias could be quite large. It appears, however, that when the exposure of each population is estimated by a single common measure (e.g. area air pollution), as would be the most useful for the quantitative estimation of risk, the bias would usually be towards the null value (Brenner *et al.*, 1992). This case, however, has not been rigorously worked out.

4.3.2 Measurement error in confounding

Error in the measurement of a confounding variable influences the results of statistical control of the effects of this variable on the dose–response relationship of the primary exposure variable with disease.

4.3.2.1 Analytical epidemiological studies

When the confounder is measured with non-differential error and the exposure is measured perfectly, the effects of the confounder will be incompletely controlled and the effects of the exposure, independent of the confounder, will appear greater or smaller than they really are, depending on the direction of confounding. When the exposure is measured with non-differential error and the confounder is measured perfectly, the adjusted disease–exposure association may be biased towards the null value to an even greater degree than the crude association. When both are measured with error, the adjusted disease–exposure association may, with reference to the true disease–exposure association, be biased towards or away from the null.

4.3.2.2 Ecological studies

Confounding in ecological studies is more complex than in analytical studies, its effects are potentially much greater, and its control more difficult (Greenland & Morgenstern, 1989; Greenland, 1992). The effects of error in the measurement of confounders on the results of ecological studies has not been described.

4.3.3 Example of effects of exposure measurement error on QEP

Much work has been carried out to evaluate and take into account sources of systematic and random error in the radiation dosimetry of the atomic bomb survivors. Around 1986, a new dosimetry system, "DS86", was introduced to replace the dose estimates, referred to as "T65D", which had previously been used for risk estimation in studies of atomic bomb survivors. Correlations of doses with chromosome aberration data indicate that DS86 doses are more accurate than T65D doses, particularly for some groups of survivors with complicated shielding patterns. DS86 doses were estimated for a number of specific target organs and are also more accurate than T65D doses, which were based only on the application of a set of organ transfer factors. Overall, DS86 neutron doses were less than the corresponding T65D ones. DS86 doses from γ-rays, however, were generally larger than the corresponding T65D ones, the difference increasing with distance from the epicentre and varying with the characteristics of shielding, irrespective of whether the survivor was a child or an adult at the time of the bombing and the target organ.

The introduction of DS86 in risk estimates coincided with a change in the risk models used to estimate and predict risk from data on atomic bomb survivors and with the analysis of extended follow-up data. The results of the life-span study follow-ups (Preston *et al.*, 1987; Shimizu *et al.*, 1990) are therefore not strictly comparable. However, Preston & Pierce (1988) carried out analyses to evaluate the impact of the change in dosimetry on risk estimates. Their results indicate that the improvement in the accuracy of the dose estimate resulted in a 75–85% increase in estimates of radiation-induced mortality from leukaemia and from all other cancers.

Following this, methods were developed (Pierce *et al.*, 1990, 1992) to adjust risk estimates for the biases resulting from random errors in individual dose estimates. The magnitude of the random errors varied with dose, being greatest for survivors with doses above 4 Gy. It was concluded that random errors of 30–40% can lead to underestimation of the risk of solid cancers by 7–11% and of leukaemia by 4–7%.

PM3006483535

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

The adequacy of neutron dosimetry, particularly in Hiroshima, has recently been questioned (Straume *et al.*, 1992). It appears that DS86 does not fully account for the neutrons probably present at Hiroshima, particularly at the greater distances, and additional work will need to be carried out to improve the dosimetry.

4.3.4 Conclusions

Error in measurement of exposure can have substantial and largely unpredictable effects on the direction and shape of dose–response relationships. The consequences of exposure measurement error for QEP are potentially serious, therefore, unless the effects of error can be controlled or error can be prevented from occurring in the first place.

4.4 Prevention of exposure measurement error

To the extent that an economically and logistically feasible way exists to achieve the prevention of errors in measurements of the primary exposure and potential confounders, this is the best way to control the effects of measurement error and therefore to minimize its effects on QEP. There are two general approaches to the prevention of measurement error: (i) quality in the design of measurement instruments; and (ii) quality control in making and using measurements.

4.4.1 Quality in the design of measurement instruments

Instruments for measuring exposure to carcinogens for the purposes of QEP should be designed to make the right measurements accurately. There are few general prescriptions for accuracy at the design stage as its determinants vary from one kind of measurement instrument to another. The removal of human subjectivity, however, is a desirable characteristic of all measurement processes. In the administration of a questionnaire by interview, for example, this will require the use of carefully trained interviewers, a highly structured questionnaire, and keeping the interviewers ignorant, in the specific situation, of whether they are interviewing a person with or without disease or with or without exposure. In laboratory measurements on environmental or biological samples, it will require the processing of all samples without those concerned knowing whether they come from subjects with or without disease.

Some specific procedures for ensuring high quality in the design of instruments for the field collection of epidemiological data, including the collection of biological samples, are summarized in Table 4.2.

Table 4.2 Quality in the design of measurement procedures for the collection of data on exposure in epidemiological studies[a]

Design of forms:
Include all items needed to compute dose, timing of exposure, etc.
Include adequate subject identifiers—at least an identity number and a check digit or alphabetic code on all forms
Use separate forms for each method of exposure measurement
Make instructions clear and data collection items unambiguous
Use different typefaces for instructions, data collection items and responses
Provide mutually exclusive and exhaustive response categories for closed-ended items
Make forms self-coding for simple items, e.g. data collector circles a number corresponding to the appropriate response category
Make response codes consistent within and across forms, e.g. 1 = yes, 2 = no
Provide for coding without loss of information, i.e. do not design forms so that continuous data are categorized at the coding stage

Table 4.2 (Contd) Quality

Design of forms (contd):
Do not require computation by d
Design forms for direct entry of d
Design of specimen collection:
Design procedures for specimen laboratory procedures
Study procedures manual:
Always have a study procedures
Include at least the following in th
• *Description of the study in ge*
• *Sample selection, recruitment*
• *Data forms*
• *General methods of data colle*
• *Item by item clarification of qu*
• *Detailed procedures for the c*
• *Editing procedures*
• *Coding instructions for items*
• *Codebooks*
Update manual and distribute up
Pretesting instruments:
Have instruments reviewed by oth
Pretest instruments on samples of
Train data collectors and pretest in
Identify problems through feedback observing interviews, reabstracting
Review frequencies of responses
Modify instrument
Pretesting instruments:
Discuss importance of complete an
Review study manual
Practise data collection
Monitor initial data collection by ea
Resolve problems

[a] Source: Armstrong *et al.*, (1992)

PM3006483536

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Table 4.2 (Contd) Quality in the design of measurement procedures for the collection of data on exposure in epidemiological studies[a]

Design of forms (contd):
Do not require computation by data collectors, rather make provision on the form for entry of raw data into the computer
Design forms for direct entry of data into the computer
Design of specimen collection:
Design procedures for specimen collection, identification, transport and storage in consultation with specialists in the laboratory procedures
Study procedures manual:
Always have a study procedures manual
Include at least the following in the study procedures manual:
• *Description of the study in general terms*
• *Sample selection, recruitment and tracking procedures*
• *Data forms*
• *General methods of data collection*
• *Item by item clarification of questions and responses on forms and questionnaire, including special cases*
• *Detailed procedures for the collection of biological samples*
• *Editing procedures*
• *Coding instructions for items not self-coded on form*
• *Codebooks*
Update manual and distribute updated pages whenever procedural changes are made
Pretesting instruments:
Have instruments reviewed by other researchers
Pretest instruments on samples of convenience
Train data collectors and pretest instruments on samples similar to study subjects
Identify problems through feedback from pretested subjects and data collectors and by monitoring data collection (e.g. observing interviews, reabstracting records) and make appropriate changes as early as possible
Review frequencies of responses to identify items with little variation in responses
Modify instrument
Pretesting instruments:
Discuss importance of complete and accurate data
Review study manual
Practise data collection
Monitor initial data collection by each data collector
Resolve problems

[a] Source: Armstrong *et al.*, (1992)

PM3006483537

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

4.4.2 Quality control in making and using measurements

The most important steps in quality control in making and using measurements are ensuring that the measurement procedures are written down, that specific quality-control measures form part of these procedures, and that all staff involved in making measurements are adequately trained. Table 4.3 summarizes specific quality-control procedures in the field collection, processing and statistical analysis of epidemiological data. Some major aspects of quality control in the laboratory analysis of biological or environmental samples collected as part of an epidemiological study are also listed. More comprehensive coverage of quality control in analytical laboratories can be found in Westgard & Klee (1986) and Copeland (1989).

Table 4.3 Quality control in making and using exposure measurements in epidemiological studies[a]

QUALITY CONTROL DURING DATA COLLECTION
Supervision of data collectors:
Assign cases and controls in a case–control study (or exposed and unexposed subjects, where this is known in advance, in a cohort study) in the same proportion to each data collector
Maintain ignorance of data collectors to case–control (exposed-unexposed) status of subjects, as far as possible
Replicate some proportion of data collection (e.g. 10% of subjects) to identify fictitious data, items with poor reliability, data collectors with errors on certain items, etc.
Compare the distribution of study variables among data collectors
Compare distributions of study variables over time
Address problems identified through monitoring immediately with the relevant data collector(s)
Conduct staff meetings for retraining, discussion of problems and motivation
Editing and coding:
Have data collectors edit data forms immediately to clarify responses and check for missing items
Have supervisor perform a second edit soon after data collection to check for missing items, inadmissible codes, inconsistencies among responses, illegible responses, etc.
Have editor code open-ended questions and query those inadequately answered
Correct errors by call back to subjects (or check-backs to records)
Have one staff member maintain an editors' log to ensure consistency of recording and coding of unanticipated responses, and to record comments and responses coded as "other"
Handling and analysis of biological samples:
Ensure as far as possible that samples are collected at around the same time from matched subjects with and without disease, or numerically balanced samples of subjects with and without disease, so that their storage and analysis times can be similar
Record time and conditions of collection, initial storage, transport, entry into secondary storage, and removal for laboratory processing, review periodically and investigate reasons for departure from agreed standards
Reasons should be given by the laboratory for rejection of samples as unacceptable and investigated in case they point to errors in collection, storage and transport procedures that can be corrected

Table 4.3 (Contd) Quali

Handling and analysis of biological
Ensure, as far as possible, that match jects with and without disease, are in
Undertake all laboratory analyses with
Implement local, within-batch quality c Optimally, participate also in a progra able, assay blind duplicates of some s
QUALITY CONTROL DURING DATA PROCESS
Key entry:
Create a codebook with format and co
Enter data contemporaneously with da
Double enter (verify) all data
Edit data by computer by performance
Correct errors and feed back findings
Creation of new variables:
Check and recheck the programming c
Check the correctness of new variabl ably possible
Review distributions of original and cr
Create a codebook with detailed descr code used to create them

[a] Source: Armstrong *et al.*, (1992)

4.4.3 Other factors

Much good advice can be given vention of error in measurement o it is self-evident that at least some dures proposed above will lead to How free from error measurement: when all of these procedures are depend very much on the exposu sured and the options available fo ment.

Some exposures, such as the sm rettes, are comparatively easy to me so because, traditionally, cigarette been a habit commenced in youth

PM3006483538

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Table 4.3 (Contd) Quality control in making and using exposure measurements in epidemiological studies[a]

Handling and analysis of biological samples (contd):
Ensure, as far as possible, that matched subjects with and without disease, or numerically balanced samples of subjects with and without disease, are included in each batch of laboratory analyses
Undertake all laboratory analyses without knowledge of which samples come from subjects with and without disease
Implement local, within-batch quality control and standardization procedures according to best laboratory practise. Optimally, participate also in a programme of external standardization of the analyses. If reference sample are not available, assay blind duplicates of some samples in the same batch and in different batches
QUALITY CONTROL DURING DATA PROCESSING
Key entry:
Create a codebook with format and code of "raw" data
Enter data contemporaneously with data collection
Double enter (verify) all data
Edit data by computer by performance of range and logic checks contemporaneously with data entry
Correct errors and feed back findings of relevance to data collection
Creation of new variables:
Check and recheck the programming code used to create new variable
Check the correctness of new variables by manual computation from a sample of original records, whenever reasonably possible
Review distributions of original and created variables
Create a codebook with detailed descriptions of new variables created including original variables and programming code used to create them

[a] Source: Armstrong *et al.* (1992)

4.4.3 Other factors

Much good advice can be given about the prevention of error in measurement of exposure and it is self-evident that at least some of the procedures proposed above will lead to its reduction. How free from error measurements may become when all of these procedures are adopted will depend very much on the exposure being measured and the options available for its measurement.

Some exposures, such as the smoking of cigarettes, are comparatively easy to measure. This is so because, traditionally, cigarette smoking has been a habit commenced in youth and indulged in more or less continuously until the habituee relinquishes it, usually in middle life, or dies. Its adoption and relinquishment are important events and are therefore easily remembered. Moreover, the number of cigarettes smoked per day is reasonably constant and also easily remembered because the smoker must regularly purchase fresh supplies. However, while empirical data indicate that cigarette consumption is recalled well if constant, it may be recalled with bias if consumption has changed (Persson & Norell, 1989). Prospectively collected data on cigarette consumption have been used to produce plausible QEPs (see, e.g. Doll & Peto, 1978; Doll *et al.*, 1994).

PM3006483539

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

However, the measurement of cigarette smoking and, in particular, the consequential exposure to carcinogens, has become more complicated over the past 30 years because of changes resulting from concerns about the health effects of smoking. The use of filter cigarettes has become almost universal in industrialized countries, cigarettes that deliver smaller amounts of tar have been designed and their use has increased substantially. Thus, to estimate the intake of tar from tobacco smoke retrospectively now requires that smokers recall when they took up the use of filter cigarettes and the brands that they smoke—a rather more demanding task. Smokers may now also tend to underestimate their cigarette consumption because of social disapproval of their habit. While none of these problems is insurmountable, it is now more difficult to obtain an accurate measure of intake of tobacco smoke carcinogens (a complex mixture for which simple cigarette consumption was a proxy) than it used to be.

Some other exposures are extremely difficult to measure no matter how carefully they are approached. For example, dietary intake of sodium cannot be measured accurately by any form of dietary assessment except the observation, weighing and sampling, for analysis, of all food eaten in some defined period of time. Much simpler than this, although by no means easy, is the measurement of urinary sodium excretion which, if carried out over 14 periods of 24 h, provides a reasonably accurate reflection of present sodium intake (Liu *et al.*, 1979). Neither of these approaches can provide a measure of anything other than present intake of sodium unless repeated. Thus accurate retrospective measurement of individual intake of sodium for, say, a case–control study of cancer of the stomach (with which it may be associated) must be considered to be impossible unless it can be assumed that dietary sodium intake is reasonably constant over time. This assumption is unlikely to be true.

Between these two extremes lie most other exposures of interest in quantitative estimation and prediction of risk of cancer. All are measured with some error. Thus steps must be taken to control the effects of the error or it will influence the results of QEP.

4.5 Control of the effects of exposure measurement error

Issues relating to the control of the effects of exposure measurement error have been discussed by Willett (1989). There are three possible approaches to control when measurement error has not been prevented (Greenland & Robins, 1985; Armstrong *et al.*, 1992): (i) the use of multiple measures of exposure; (ii) the adjustment of study results for the effects of measurement error; and (iii) in some circumstances, adjustment in the analysis for a covariate which is related in some way to the occurrence of error in the exposure measurement. In principle, each of these approaches can produce a more accurate representation of the dose–response relationship and therefore allow more accurate QEP.

4.5.1 Multiple measures of exposure

The average or some other combination of the information in multiple measures of the same exposure for each individual in an analytical study (or population in an ecological study) is, in principle, a more accurate measure of exposure than a single measurement because of reduction in within-subject variability (Marshall, 1989; Brenner & Blettner, 1993). The multiple measures may be either repeated measures using the same (e.g. repeated collection of dietary diaries over a one-year period) or different instruments (e.g. a combination of measurements from a food frequency questionnaire and a period of collection of dietary diaries). There may be no advantage and may even be a disadvantage in the latter approach if one of the two instruments is substantially less accurate than the other. The average of the two measurements may then be less accurate than the measurement with the more accurate instrument.

The use of multiple measures of exposure is most unlikely to fully control the effects of measurement error. In the case of repeated applications of the same instrument, for example, no correction will be obtained for bias in measurements or in error which differs between subjects with and without disease. While some correction of both may be achieved if two different measurement instruments are used, full correction will only be possible if both instruments were perfect in the first place o opposite biases, which is highl

4.5.2 Adjustment of study result measurement error

Table 4.4, taken from Armsi gives equations expressing regression coefficients or odds

Table 4.4. Equations for measure of

Equation[b]
Continuous exposure, dichotomous
$\beta_T = 1 - \left[\frac{\bar{x}_D - b_N}{\mu_{X_D} - \mu_{X_N}}\right]\frac{\beta_O}{\rho^2_{TX}}$
$\beta_T = \beta_O / \rho^2_{TX}$
$OR_T = OR_O^{1/\rho^2_{TX}}$
Continuous exposure, continuous ou
$\beta_T = \beta_O / \rho^2_{TX}$
Dichotomous exposure, dichotomou
$OR_T = \frac{P_D(1-P_N)}{P_N(1-P_D)}$
where $P_D = (\rho_D - 1 + \text{spec}_D)/(\text{se}$
and $P_N = (\rho_N - 1 + \text{spec}_N)/(\text{sens}$

[a] Source: Armstrong *et al.* (1992).

[b] β_T and β_O are true and observable log the observable measurements in disea nents in diseased and non-diseased e cient), ρ_D and ρ_N are the observable pr the observable measurements in disea surements in diseased and non-disease

In practice, the position is muc First, there may be no satisfactory the validity of the exposure meas made because no perfect meas exists. A comparison of the meth

PM3006483540

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

effects of exposure mea-

the control of the effects of
ent error have been dis-
89). There are three possi-
ntrol when measurement
prevented (Greenland &
ng *et al.*, 1992): (i) the use
of exposure; (ii) the adjust-
for the effects of measure-
in some circumstances,
lysis for a covariate which
to the occurrence of error
surement. In principle,
hes can produce a more
n of the dose–response
fore allow more accurate

s of exposure
e other combination of
ultiple measures of the
individual in an analyti-
in an ecological study)
e accurate measure of
neasurement because of
n-subject variability
& Blettner, 1993). The
be either repeated mea-
. repeated collection of
-year period) or differ-
mbination of measure-
ncy questionnaire and
dietary diaries). There
l may even be a disad-
oach if one of the two
lly less accurate than
he two measurements
han the measurement
trument.
asures of exposure is
rol the effects of mea-
e of repeated applica-
ent, for example, no
l for bias in measure-
fers between subjects
While some correc-
eved if two different
are used, full correc-
if both instruments

were perfect in the first place or had had exactly opposite biases, which is highly unlikely.

4.5.2 Adjustment of study results for the effects of measurement error

Table 4.4, taken from Armstrong *et al.* (1992), gives equations expressing the true logistic regression coefficients or odds ratios in terms of their observable values and exposure measurement error. In principle, these equations can be used to estimate the true values given the observed values and estimates of the error expressed in appropriate terms. Estimates of the error can be obtained by means of a validity study in a sample of the subjects in the epidemiological study.

Table 4.4. Equations for the true measure of association as a function of the observable measure of association and the exposure measurement error[a]

Equation[b]	Type of error
Continuous exposure, dichotomous outcome:	
$\beta_T = 1 - \left[\frac{b_D - b_N}{\mu_{X_D} - \mu_{X_N}}\right] \frac{\beta_O}{\rho^2_{TX}}$	Differential
$\beta_T = \beta_O / \rho^2_{TX}$	Non-differential
$OR_T = OR_O^{1/\rho^2_{TX}}$	Non-differential
Continuous exposure, continuous outcome:	
$\beta_T = \beta_O / \rho^2_{TX}$	Non-differential
Dichotomous exposure, dichotomous outcome:	
$OR_T = \frac{P_D(1 - P_N)}{P_N(1 - P_D)}$	Differential or non-differential
where $P_D = (p_D - 1 + \text{spec}_D)/(\text{sens}_D + \text{spec}_D - 1)$	
and $P_N = (p_N - 1 + \text{spec}_N)/(\text{sens}_N + \text{spec}_N - 1)$	

[a] Source: Armstrong *et al.* (1992).

[b] β_T and β_O are true and observable logistic regression coefficients, OR_T and OR_O are true and observable odds ratios, b_D and b_N are bias in the observable measurements in diseased (D) and non-diseased (N) subjects, μ_{X_D} and μ_{X_N} are expectations of the observable measurements in diseased and non-diseased subjects, ρ_{TX} is the correlation between the true and observable measurements (the validity coefficient), p_D and p_N are the observable proportions exposed in diseased and non-diseased subjects, spec_D and spec_N are the specificities of the observable measurements in diseased and non-diseased subjects, and sens_D and sens_N are the sensitivities of the observable measurements in diseased and non-diseased subjects.

In practice, the position is much more complex. First, there may be no satisfactory way of measuring the validity of the exposure measurements actually made because no perfect measure of exposure exists. A comparison of the method actually used with what is presumed to be a more accurate method may be made to provide an upper limit estimate of the validity coefficient, ρ_{TX}, in the case of a continuous measurement of exposure, and therefore a probably conservative correction for the

PM3006483541

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

effects of measurement error. Second, the equations in Table 4.4 assume, among other things, that the extent of measurement error is uncorrelated with the value of the true exposure. If this and other assumptions are not true, the adjustment will not be valid. Third, allowance needs to be made for sampling error in the observable disease–exposure association and the measurements of error. Finally, information on the multivariate measurement error structure of the primary exposure variable and the confounding variables is required if an adjustment for error in the measurement of all relevant variables is to be attempted.

While work has been done on a number of the problems listed above, and procedures are available that take account of some of them (Armstrong *et al.*, 1992; Thomas *et al.*, 1993), it will in most cases be difficult to overcome the difficulty that a source of true measurements against which a sample of the measurements actually obtained can be compared is lacking (Wacholder *et al.*, 1993). Therefore, while dose–response relationships adjusted for the effects of measurement error may be nearer the truth than the unadjusted relationships, they are unlikely to be close to perfect. Moreover, because the sensitivity and specificity or the validity coefficient will often have been derived from a small sample of subjects, the sampling error in any adjusted dose–response relationship or prediction of risk will be correspondingly large.

An example of the adjustment of a continuous dose–response relationship with reference to a validation study carried out in the same population of subjects is shown in Table 4.5, based on the US Nurses' Health Study (Rosner *et al.*, 1989). This examines the effect of using a comparison of calorie-adjusted, daily intake of saturated fat, measured both by food frequency questionnaire and by multiple daily records of weighed dietary intake in a sample of 173 nurses, to adjust the estimate of relative risk per 10 g increase in calorie-adjusted saturated fat intake based on the food frequency questionnaire. As would be expected, the adjusted relative risks are further from the null value (1.0) than the unadjusted estimates, and their 95% confidence intervals are wider because of the component of sampling error in the coefficient of regression of the the two estimates of adjusted saturated fat intake.

While the bias in the unadjusted estimate, in this example is apparently not large the adjusted estimate is open to question. First, while the recording of weighed dietary intake is probably the most accurate method of estimating nutrient intake, it is not perfect. For example, the close observation of diet that it entails is likely to change dietary intake in unpredictable ways. Second, while the 28 days of recording were spread throughout one year, they could only relate to diet in the recent past. The food frequency questionnaire, on the other hand, co
related to diet over a much l
however, is likely to have b
sure of the period of life in w
may be most important in b
namely, childhood (de Waar

4.5.3 Control of a covariate
ment error

It is common in epidem
either in the design or an
ensure comparability in m
jects with and without dis
when multiple interview
case–control study, each c
matched controls may be
same interviewer and at abo
equalize, between cases and
error and any error in respon
with time. Similarly, when p
necessary for some subjects
died before an interview c
proxy may also be used for t
or the analysis may be strat
subject or a proxy was inter
1988). Such strategies will s
in the effect measures. As a
classification is non-differen

Table 4.5 Use of results of a validation study to adjust for error in measurement of exposure: relationship of risk of breast cancer to calorie-adjusted saturated fat intake in a cohort of US nurses

Validation study		Main cohort study	
Days of diet record	Estimate of regression coefficient (SE)[b]	Observed relative risk (95% CI)[c]	Adjusted relative risk (95% CI)[c]
28	0.468 (0.048)	0.92 (0.80–1.05)	0.83 (0.61–1.12)
2	0.540 (0.090)	0.92 (0.80–1.05)	0.85 (0.65–1.11)

[a] Based on data from Rosner et al. (1989).
[b] Coefficient of linear regression of estimated calorie-adjusted saturated fat intake from diet records on estimate from food frequency questionnaire. SE = standard error.
[c] Relative risk per 10g increase in saturated fat intake. CI = confidence interval.

**Table 4.6 Effect of adj
of proxy**

Risks	Current sm	
		Non-smoke
True relative risks	1.0	
Crude observed relative risks	1.0	
Observed relative risks adjusted for respondent	1.0	

[a] Based on data from Walker et al.
smoking habits in which 50% of e

PM3006483542

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

naire, on the other hand, could, in principle, have related to diet over a much longer period. Neither, however, is likely to have been an accurate measure of the period of life in which dietary variables may be most important in breast cancer aetiology, namely, childhood (de Waard, 1986).

4.5.3 Control of a covariate related to measurement error

It is common in epidemiology to take steps either in the design or analysis of a study to ensure comparability in measurements on subjects with and without disease. For example, when multiple interviewers are used in a case–control study, each case and his or her matched controls may be interviewed by the same interviewer and at about the same time to equalize, between cases and controls, interviewer error and any error in response that may correlate with time. Similarly, when proxy respondents are necessary for some subjects (usually because they died before an interview could be obtained), a proxy may also be used for the matched controls, or the analysis may be stratified on whether the subject or a proxy was interviewed (Walker *et al.*, 1988). Such strategies will sometimes reduce bias in the effect measures. As a general rule, if misclassification is non-differential within levels of a covariate but varies across those levels, the validity of the effect measure will be improved by control of the covariate. In other circumstances, control of a covariate related to measurement error may increase bias in the effect measure (Greenland & Robins, 1985).

Walker *et al.* (1988) have given a hypothetical example of the effect of control of respondent status—self or proxy—in a case–control study of the relationship between smoking and a disease. They assumed that proxy responses were obtained for 50% of cases and 10% of controls, and used the matrix of proxy-reported versus self-reported smoking habits obtained by Rogot & Reid (1975). The results are summarized in Table 4.6. The effect of the differential measurement error introduced by the two different sources of exposure measurements is to slightly increase the relative risk of disease in smokers of more than one pack per day and to alter the shape of the dose–response relationship at the bottom of the distribution of cigarette use. Adjustment for respondent type corrects both of these biases, but increases slightly the width of the 95% confidence intervals about the relative risk estimates because of maldistribution of controls between the two groups of cases of equal size in the stratified analysis.

Table 4.6 Effect of adjustment for a source of exposure measurement error [illegible] of proxy respondents, on bias due to the misclassification[a]

Risks	Current smoking habits								
	Non-smoker	Occasional smoker		< 1 pack per day		1 pack per day		> 1 pack per day	
		RR	95% CI	RR	95% CI	RR	95% CI	RR	95% CI
True relative risks	1.0	1.4	0.5–3.5	2.0	1.4–2.9	3.0	2.1–4.3	3.5	2.3–5.3
Crude observed relative risks	1.0	2.2	1.0–4.7	1.7	1.2–2.5	2.9	2.1–4.2	3.7	2.5–5.6
Observed relative risks adjusted for respondent	1.0	1.5	0.6–3.6	2.0	1.3–3.1	2.9	2.0–4.3	3.3	2.1–5.2

[a] Based on data from Walker et al. (1988). Hypothetical example of a case–control study of the relationship between a disease and current smoking habits in which 50% of exposure measurements in cases came from proxy respondents but only 10% in controls.

PM3006483543

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

4.5.4 Resources for the control of exposure measurement error

Strategies to control the effects of exposure measurement error without actually preventing it can reduce the effects of both bias and random error on the dose–response relationship and therefore lead to more accurate QEP. None of these strategies is perfect, however, and it is unlikely that use of any or all of them will eliminate the effects of error entirely. In addition, two of them, namely the use of multiple measurements of exposure and adjustment of results with reference to a validity or reliability study, are expensive and may be difficult in practice or, in some cases (e.g. most historical cohort studies), impossible.

The allocation of resources to preventing or controlling the effects of exposure measurement error may well be cost-effective given the consequences of measurement error on the numbers of subjects needed to detect an effect or estimate it with some given precision. Freedman *et al.*, (1990) estimated the effects of error in the measurement of dietary fat intake on the numbers of subjects required in a cohort study to detect, with 90% power at $p<0.05$, relative risks of colorectal cancer of up to 2.3 over five levels of proportion of dietary calories derived from fat. If the correlation in the cohort between the measured and true proportion of dietary calories derived from fat was 0.65, they showed that a cohort of 141 000 subjects with an expected incidence of colorectal cancer of 200 per 100 000 person–years followed for five years would be required. The number required if the *true* exposure could be measured was nearly 10-fold less, namely 16 000.

To what should resources for the prevention or control of the effects of measurement error be allocated? In general, it appears that, when the proposed measure of exposure is reasonably accurate, a validation study on a sample of the study population, and use of its results to correct the risk estimates obtained in the whole population, is the most cost-effective approach. When the proposed measure is very inaccurate, it may be more efficient simply to use the criterion measurement proposed for the validation study as the measure to be applied in the whole population (Greenland, 1988; Spiegelman & Gray, 1991). Where there is no criterion measure, it may in some circumstances be more efficient to repeat the planned measure in all subjects and reduce the total sample size than to measure only once (Phillips & Davey Smith, 1993).

4.6 Error in the measurement of outcome

In cancer epidemiology, the outcome of interest is the occurrence of cancer in the subjects under study. However, although easy to conceptualize, occurrence of cancer is a difficult event to measure. As a result, the event actually measured may be more or less closely related to it.

The concept of occurrence of cancer is ambiguous. For example, how one considers a case of clinically silent cancer detected at necropsy will depend on the context of the study: it will be counted as a case in a study on the validation of death certificates by way of necropsy records but will be ignored in a study based solely on death certificates, and one could argue about its inclusion in a population-based case–control study. Even the most straightforward definition of a case of incident cancer, based on medical records reporting the results of diagnostic procedures and clinical reasoning, may be subject to ambiguity since medical procedures and diagnostic criteria differ among medical centres and among patients referred to the same centre (for a discussion on the comparability of data in cancer registries, see Parkin & Muir, 1992).

The most widely used outcome variables in cancer epidemiology, namely death from cancer, diagnosis of cancer, diagnosis of preneoplastic lesions, and molecular markers of early carcinogenic effect, can be seen as consecutive steps in a process corresponding to the putative natural history of the disease (for a review, see MacMahon & Pugh, 1970; Hulka *et al.*, 1990; and section 4.2.3).

4.6.1 Death from cancer

A fundamental source of error in outcome ascertainment is incorrect specification of vital status. In particular, deaths may be missed when death records are incomplete. For example, in studies involving record linkage with mortality records, vital status may be incorrectly specified when personal identifiers are insufficient to ensure that no linkage errors occur. Linkage errors lead to bias and additional uncertainty in commonly used statistics such as standardized

mortality ratios (SMRs) and
sion coefficients, with an exces
leading to upward (downwar
mated SMRs.

Information on death fro
mainly on death certificates,
legal reasons in many cou
sources of information on deat
include medical records and in
of-kin (see, e.g. Chen *et al.*, 19
ological studies of occupatio
based on mortality data from

The main advantage of the
cates is their availability in ma
for relatively long periods, allo
low-up of large groups of expo
derivation of reference rates i
studies. Among the proble
their use are limitations in the
death certificates resulting, e.
in survival in different popu
report cancer in patients dying
and limitations in the accura
themselves, including, in parti
reported cause of death and in
In a study of 1243 death certi
cancer, which were coded ind
countries, the proportion of
with cancer as the underlying
ied from 88% to 96% (Percy &

4.6.2 Incidence of cancer

Data on cancer incidence
from established systems such
reporting routinely all cases oc
ular population (see, e.g. Park
from ad hoc surveys of medic
(e.g. in a hospital-based case-c
use of data from high-quality c
several theoretical advantages,
control of the quality and com
tration and minimizing diagn
tion (Parkin & Muir, 1992). H
cancer registration is often u
cancer registration covers less
individuals worldwide (Park
Studies using data on cancer
from ad hoc surveys are sub
problems because of their pot
ness, which may lead to bias w

PM3006483544

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

more efficient to repeat
n all subjects and reduce
ıan to measure only once
ı, 1993).

rement of outcome

gy, the outcome of inter-
of cancer in the subjects
although easy to concep-
ncer is a difficult event to
e event actually measured
ely related to it.
ence of cancer is ambigu-
one considers a case of
letected at necropsy will
of the study: it will be
udy on the validation of
of necropsy records but
ly based solely on death
ld argue about its inclu-
sed case–control study.
ward definition of a case
ed on medical records
agnostic procedures and
e subject to ambiguity
and diagnostic criteria
tres and among patients
tre (for a discussion on
in cancer registries, see

l outcome variables in
ely death from cancer,
nosis of preneoplastic
arkers of early carcino-
s consecutive steps in a
he putative natural his-
view, see MacMahon &
990; and section 4.2.3).

of error in outcome
specification of vital
s may be missed when
lete. For example, in
inkage with mortality
e incorrectly specified
s are insufficient to
rrors occur. Linkage
litional uncertainty in
such as standardized

mortality ratios (SMRs) and relative risk regression coefficients, with an excess (deficit) of deaths leading to upward (downward) bias in the estimated SMRs.

Information on death from cancer is based mainly on death certificates, which are issued for legal reasons in many countries. Additional sources of information on death from cancer may include medical records and interviews with next-of-kin (see, e.g. Chen *et al.*, 1990). Most epidemiological studies of occupational cancer risk are based on mortality data from death certificates.

The main advantage of the use of death certificates is their availability in many populations and for relatively long periods, allowing long-term follow-up of large groups of exposed people and easy derivation of reference rates in population-based studies. Among the problems associated with their use are limitations in the data obtained from death certificates resulting, e.g. from differences in survival in different populations, failure to report cancer in patients dying from other causes, and limitations in the accuracy of the certificates themselves, including, in particular, errors in the reported cause of death and in the coding process. In a study of 1243 death certificates mentioning cancer, which were coded independently in nine countries, the proportion of certificates coded with cancer as the underlying cause of death varied from 88% to 96% (Percy & Muir, 1989).

4.6.2 Incidence of cancer

Data on cancer incidence are usually derived from established systems such as cancer registries reporting routinely all cases occurring in a particular population (see, e.g. Parkin *et al.*, 1992), or from ad hoc surveys of medical or other records (e.g. in a hospital-based case–control study). The use of data from high-quality cancer registries has several theoretical advantages, such as ensuring control of the quality and completeness of registration and minimizing diagnostic misclassification (Parkin & Muir, 1992). However, quality of cancer registration is often unsatisfactory, and cancer registration covers less than 400 million individuals worldwide (Parkin *et al.*, 1992). Studies using data on cancer incidence derived from ad hoc surveys are subject to additional problems because of their potential incompleteness, which may lead to bias with respect to the assumed population base, and because the information derived from sources other than medical records may not be correct (see, e.g. Steenland & Schnorr (1988) on diagnoses reported by next of kin).

One additional problem related to the use of data on cancer incidence is the fact that the degree of medical surveillance may vary among populations and groups, and thus lead to variable ascertainment of the true underlying incidence. Examples of such surveillance bias may be found in cancer patients followed up for incidence of secondary leukaemia (Kaldor *et al.*, 1990) and occupational groups undergoing periodic medical examinations (Hiatt *et al.*, 1993).

Whether the cases of cancer included in a study are obtained from mortality statistics, from a cancer registry or from an ad hoc survey (e.g. from the of pathology department of a hospital), measurement error can be introduced in the diagnosis of the neoplasm. In particular, there can be a diagnostic error when the true disease is not diagnosed, and a histopathological error when the measured histopathological class is not the true class. It should be noted that histopathological error is not necessarily less serious than diagnostic error: the inclusion of benign lesions in a series of malignant tumours, and similarly the exclusion of malignant lesions misdiagnosed as benign tumours, will have very important implications in a study on the quantification of the effect of a risk factor.

These errors can either be systematic or non-systematic, as discussed earlier for exposure measurement errors. Thus a systematic histopathological error exists, e.g. when a pathologist has a consistent tendency to call a benign melanocytic naevus a level I melanoma.

The major difficulty in assessing diagnostic or histopathological validity is the lack of a gold standard of the characteristics being measured. In other words, there is no way to determine the absolute truth with respect to the disease at the level of organ, tissue and cell type. For this reason, measurements of diagnostic and histopathological error have usually focused on reliability rather than validity, and have measured intraobserver and interobserver reliability.

Validity and reliability are related in the sense that valid measurements are also reliable.

PM3006483545

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Reliable measurements, however, are not necessarily valid, and the only certain conclusion is that an unreliable measurement is also invalid. Intraobserver reliability tends to be greater than the corresponding interobserver reliability because the single observer will repeat his or her own systematic error from observation to observation. Thus, the comparison of repeated measurements by one observer will be free from the effects of systematic error, whereas the often partly different systematic errors of different observers will be included in the assessment of interobserver agreement. Because systematic error is an important component of outcome measurement error, a measurement of interobserver reliability will be a better guide to validity (or at least to its upper limit) than a measurement of intraobserver reliability.

4.6.3 Preneoplastic lesions

Although there are many examples of preneoplastic lesions identified and measured in human populations, more important from the point of view of quantitative associations are those in which the preneoplastic lesions are used as surrogates for the cancer itself. In such cases, a quantitative estimation of risk can be made on the basis of the presence, or level, of the preneoplastic lesion. For example, melanocytic naevi are strong predictors of malignant melanoma of the skin (Armstrong & English, 1988), and genital warts and other manifestations of infection with human papilloma virus (HPV) are strongly associated with the risk of cervical cancer (Brinton, 1992).

Measurement of preneoplastic lesions, however, is also subject to random and systematic error. For example, in a study on the prevalence of melanocytic naevi in Western Australia in children assessed by two nurses, it was found that 4% of the variation in numbers of naevi of all sizes, and 8% of the variation in those of naevi of at least 2 mm in diameter was due to interobserver variation (English & Armstrong, 1994). Detection of HPV infection strongly depends on the method used, and recently developed methods based on the polymerase chain reaction have improved the sensitivity of the assay (Schiffman *et al.*, 1991) and led to a reassessment of the quantitative aspects of the association between HPV infection and risk of cervical cancer (Muñoz *et al.*, 1992).

4.6.4 Molecular markers of early carcinogenic effect

More recently, biological markers of early effects in the carcinogenic pathway have been used in studies in human populations (Perera & Weinstein, 1982). Such markers include non-specific lesions of the genome, such as sister chromatid exchanges (Wilcosky & Rynard, 1990), micronuclei (Vine, 1990) and chromosome aberrations (Schwartz, 1990), as well as more specific effects, such as activation of proto-oncogenes or inhibition of tumour-suppressor genes (Perera & Santella, 1993). In addition, biomarkers are now widely used to improve assessment of the biologically effective dose and to evaluate the role of individual susceptibility. It is clear, however, that when such methods are used, the traditional distinction between exposure and outcome becomes less clear, and the same event, such as the formation of a DNA adduct or the mutation of a relevant gene, can be considered either as a measure of exposure or as an effect.

Biomarkers of early carcinogenic effect may contribute greatly to knowledge of aetiological agents and the mechanism of tumour formation in humans (for review, see Hulka *et al.*, 1990; Perera & Santella, 1993). In particular, from the point of view of quantitative estimation and prediction of risk, they may help to improve the classification of neoplastic lesions according to the genes that have been altered and the mutations in such genes, which may be risk-factor-specific. For example, it is possible to distinguish "spontaneous" mutations in the *p53* tumour-suppressor gene from those caused by carcinogens such as benzo[*a*]pyrene or aflatoxin B_1 (Hollstein *et al.*, 1991; Jones *et al.*, 1991). There are also mutations of the *p53* gene in skin cancer and normal skin which are almost completely specific to exposure to UV radiation (Ziegler *et al.*, 1993; Nakazawa *et al*, 1994).

In addition, these biomarkers may help to identify critical steps in the carcinogenic process, and therefore to contribute to the definition of biologically relevant models of carcinogenicity to be used in quantitative human risk assessment (Perera *et al.*, 1991; Hattis & Silver, 1993). The technical variability of biomarkers has been discussed by Vineis *et al.* (1993).

4.6.5 Sources, measurement error in the measurement of c

The observer is not the only the measurement of outcom which can be classified as "pr er", and "postobserver". Table examples of these sources histopathological classificati lesions. The preobserver source tant in relation to interobserv validity in general, because they

Table 4.7. Examples

Preobserver sources:
State of preservation of tissue at th
Location of biopsy (right site, right i
Sampling of tissue for preservation
Trauma to tissue in the course of s
Preservation and fixation of tissue
Sampling of tissue section to be mo
Preparation and processing of tissu
Quality and maintenance of micro
"State of the art" of measurement of
Observer sources:
Familiarity with the "state of the art" tion
Sampling of fields of the section for
Thoroughness in making the observa mind, pressure of time)
Application of the classification rules
Postobserver sources:
Recording and transcription
Coding
Data processing
Communication
Integration of observations and class

PM3006483546

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

ers of early carcinogenic

logical markers of early genic pathway have been ian populations (Perera & markers include non-spe-iome, such as sister chro-lcosky & Rynard, 1990), 0) and chromosome aber-)), as well as more specific on of proto-oncogenes or uppressor genes (Perera & ition, biomarkers are now assessment of the biolog-d to evaluate the role of y. It is clear, however, ls are used, the tradition-exposure and outcome the same event, such as adduct or the mutation be considered either as a is an effect.

carcinogenic effect may nowledge of aetiological m of tumour formation see Hulka *et al.*, 1990; In particular, from the itative estimation and lay help to improve the ic lesions according to altered and the muta-which may be risk-le, it is possible to dis-mutations in the *p53* from those caused by o[*a*]pyrene or aflatoxin ; Jones *et al.*, 1991). of the *p53* gene in skin hich are almost com-ure to UV radiation awa *et al*, 1994). narkers may help to n the carcinogenic ontribute to the defi-vant models of car-quantitative human *al.*, 1991; Hattis & al variability of bio-ed by Vineis *et al.*

4.6.5 Sources, measurement and prevention of error in the measurement of outcome

The observer is not the only source of errors in the measurement of outcome in epidemiology, which can be classified as "preobserver", "observer", and "postobserver". Table 4.7 provides some examples of these sources of error in the histopathological classification of neoplastic lesions. The preobserver sources of error are important in relation to interobserver reliability and to validity in general, because they may contribute to measurement error by making the observer's task more difficult or increasing the likelihood of different biases between observers, and may also increase the level of variation seen between observers in a formal study of interobserver reliability (e.g. examination of different sections by different observers gives rise to the possibility of different errors in sampling from the tissue block and different quality of preparation of the mounted, stained section). A bibliography on observer variability has been published (Feinstein, 1985).

Table 4.7 Examples of sources of error in the histological classification of cancer

Preobserver sources:
State of preservation of tissue at the time of sampling
Location of biopsy (right site, right tissue)
Sampling of tissue for preservation and fixation
Trauma to tissue in the course of sampling or biopsy
Preservation and fixation of tissue
Sampling of tissue section to be mounted and stained
Preparation and processing of tissue section ready for examination
Quality and maintenance of microscope
"State of the art" of measurement of histopathological characteristics relevant to the tissue sample in question
Observer sources:
Familiarity with the "state of the art" of measurement of histopathological characteristics relevant to the tissue sample in question
Sampling of fields of the section for close examination
Thoroughness in making the observations necessary to classify the tissue (influenced by factors such as personality, state of mind, pressure of time)
Application of the classification rules implied in the "state of the art"
Postobserver sources:
Recording and transcription
Coding
Data processing
Communication
Integration of observations and classification rules into a diagnosis

PM3006483547

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Table 4.8 Requirements for the conduct of studies of interobserver reliability in histopathology

1. A representative set of sections of the cancer(s) to be studied. Ideally this set will come from the real life practice of pathology in a whole population so that the results will reflect what may actually happen in practice rather than what can be achieved in ideal circumstances
2. Two or more willing pathologists. Again, ideally these pathologists will be involved in the general practice of histopathology of *cancer rather* than be particularly specialized in the *cancer* of interest
3. Circulation of the sections to be studied, in random order and with only numerical identification, to each pathologist who examines them without knowledge of the other pathologists' opinions of them. Ideally, because of the problem of sampling of the tissue block when taking a section and variation in the quality of the subsequent processing of the section, each pathologist should see the same section
4. Standardized recording on a specially prepared form of the histopathological observations of interest. An example of a form used in a study of interobserver variation in measurement of the histopathological characteristics of malignant melanoma is given in Figure 4.3

Interobserver reliability can be measured by means of ad hoc studies involving repeated independent measurements of the outcome based on the same, or at least very similar, material. It is obvious that in some instances a correct experimental design of the reliability study is not possible, since the measurement condition cannot be replicated (e.g. diagnosis based on macroscopic necropsy evidence). Some of the requirements for the conduct of a study of interobserver reliability in the histopathology of cancer are summarized in Table 4.8.

Several statistical methods have been proposed for use in measuring interobserver reliability. The simplest and most commonly used method is to determine the proportion of cases in which two (or more) observers agree. However, this method does not take into account the proportion of cases in agreement by chance only. It can be simply shown that this proportion varies with the prevalence of the outcome being measured, increasing as the prevalence increases to very high or decreases to very low values. A simple statistic corrected for chance agreement is the kappa statistic (Fleiss, 1971), which is the ratio of the difference between the observed proportion in agreement (p_o) and the proportion of agreement that can be expected by chance (p_e) to the proportion of agreement that cannot be attributed to chance ($1-p_e$), so that:

$$\text{kappa} = \frac{p_o - p_e}{1 - p_e}$$

An advantage of the kappa statistic is that is can be applied to nominal scale variables, which are common in pathology and histology. It can be interpreted like a correlation coefficient, i.e. its value varies from 0.0 (no agreement beyond that expected by chance) to 1.0 (perfect agreement) and to -1.0 (complete disagreement). While no strict rules exist, it has been stated that values of kappa in excess of 0.75 indicate very good agreement, and values below 0.40 indicate poor agreement (Landis & Koch, 1977). More information on the kappa statistic and on its limitations can be found in specialized papers (Maclure & Willett, 1987; Mazoyer & Mary, 1987; Dunn, 1989).

Alternatives to the kappa statistic include the intraclass correlation coefficient, which is less sensitive to changes in the number of categories, and the Pearson correlation coefficient (Snedecor & Cochran, 1980).

In principle, error in outcome measurement can be prevented or reduced by dealing with the sources discussed. While every effort should be made to minimize the error in any single set of measurements, error can be reduced further if multiple measurements are made. In practice, this means the use of multiple observers. The measurement to be taken as correct when multiple observers are used may be the majority position of the observers. There is value, however, in having the observers meet after they have made their measurements, discuss their findings, and reach a consensus position on the value of the measurement (Kraemer, 1992).

1. SERIAL NUMBER:
2. TUMOUR WIDTH & BREA
   MILLIMETERS:
   (2 largest diameters)
   i. Total lesion ☐
   ii. Nodular portion ☐
3. CROSS-SECTIONAL PROFI
   1 Flat
   2 Dome
   3 Polypoid
   4 Plateau
   5 Verrucous
4. HISTOLOGY:
   (a) Premalignant intraepid without invasion
   1 HMF type
   2 SSM type
   3 Mixed
   4 Other (specify below)

   (b) Premalignant intraepid involvement adjacent to in tumour
   1 HMF type
   2 SSM type
   3 Mixed
   4 Other (specify below)
   5 None

   (c) Cell type predominatin
   1 Epithelioid
   2 Spindle
   3 Naevus cell-like
   4 Mixed

   (d) Degree of pigmentation
   1 Heavy
   2 Moderate to light
   3 None

   (e) Pre-existing or associate benign melanotic lesion
   Type.....

Figure 4.3. Form used for recording measurement of features of the histo

PM3006483548

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

MELANOMA PROJECT PATHOLOGY ABSTRACT

1. SERIAL NUMBER: ☐☐☐☐ 4

2. TUMOUR WIDTH & BREADTH IN MILLIMETERS:
(2 largest diameters)

i. Total lesion ☐☐.☐ ☐☐.☐ 10

ii. Nodular portion ☐☐.☐ ☐☐.☐ 16

3. CROSS-SECTIONAL PROFILE:
1 Flat
2 Dome
3 Polypoid
4 Plateau
5 Verrucous ☐ 17

4. HISTOLOGY:

(a) Premalignant intraepidermal lesion without invasion
1 HMF type
2 SSM type
3 Mixed
4 Other (specify below) ☐ 18

(b) Premalignant intraepidermal involvement adjacent to invasive tumour
1 HMF type
2 SSM type
3 Mixed
4 Other (specify below)
5 None ☐ 19

(c) Cell type predominating
1 Epithelioid
2 Spindle
3 Naevus cell-like
4 Mixed ☐ 20

(d) Degree of pigmentation
1 Heavy
2 Moderate to light
3 None ☐ 21

(e) Pre-existing or associated benign melanotic lesion
Type..... ☐ 23

(f) Chronic inflammatory reaction

i. Intraepidermal component
1 Severe
2 Moderate
3 Mild
4 None ☐ 24

ii. Invasive component
1 Severe
2 Moderate
3 Mild
4 None ☐ 25

(g) Evidence of regression
1 Yes
2 No ☐ 26

(h) Mitotic activity in invasive component (HPF=high power field × 400)
1 Less than 1 per 5 HPFs
2 Between 1/HPF and 1/5 HPFs
3 Greater than or equal to 1 per HPF ☐ 27

(i) Solar elastosis
1 Severe or moderate
2 Mild
3 Absent ☐ 28

(j) Ulceration
1 Yes
2 No ☐ 29

5. LEVEL OF INVASION:
1 Intraepidermal
2 Papillary dermal
3 Papillary-reticular interface
4 Reticular dermal
5 Subcutaneous fat ☐ 30

6. TUMOUR DEPTH IN MILLIMETERS:
☐☐.☐☐ 34

7. PATHOLOGIST
..... ☐ 38

Figure 4.3. Form used for recording each pathologist's observations on each case in an interobserver reliability study of the measurement of features of the histopathology of malignant melanoma. Based on Heenan *et al.* (1984).

PM3006483549

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

4.6.6 Effect of error in measurement of outcome on the dose–response relationship

The effect of outcome measurement error on the dose–response relationship is conceptually similar to the effect of exposure measurement error. When the outcome being measured is death from, or incidence of cancer, the outcome variable is invariably categorical, usually binary (i.e., presence or absence of cancer). If misclassification of a dichotomous outcome variable is independent of misclassification of exposure, the bias is always towards the null. In practice, however, misclassification of exposure and outcome might be correlated, resulting in a bias in an unpredictable direction, even with non-differential misclassification (Kristensen, 1992; Brenner *et al.*, 1993). When outcomes other than cancer are considered, there may be multiple categories, ordered, e.g. according to the severity of the disease (e.g. healthy, *in situ* carcinoma, invasive carcinoma) or reflecting multiplicity of a particular outcome in individual subjects (e.g. number of melanocytic naevi). The measurement of markers of early carcinogenic effect has added further complexity (Schulte *et al.*, 1993). While such markers are usually measured on two levels, it is now possible to identify several outcome categories based on combinations of them. For example, in a hypothetical study based on the measurement of cancer and a marker of early effect, such as activation of a proto-oncogene, four outcome categories can be considered: (1) healthy subjects without the activation; (2) healthy subjects with the activation; (3) cancer patients without the activation; and (4) cancer patients with the activation. In this example, categories 1, 2 and 4 can be seen as successive steps in the carcinogenic process, in which the presence of the early lesion is a required step. The implications of measurement error in studies based on more than two categories of outcome have not been investigated in any detail, and the effect of their misclassification on the risk estimate will be difficult to predict.

The major issue in outcome measurement error is whether the error is differential or non-differential across exposure categories. One example of differential measurement error is a cohort study in which the population under investigation (considered as exposed) is compared with a reference (e.g. national) population, considered as unexposed, where the members of the exposed cohort are more likely to be diagnosed with the cancer of interest than the reference population (see discussion on surveillance bias above).

In the case of non-differential error in measurement of outcome, by analogy with the effects of error in measurement of exposure (see section 4.3), the observed relative risk (RR) will be biased towards the null value relative to the true RR (and the dose–response curve will be flattened) provided that the probability that the measurement classifies a true case as a case is the same as, or greater than that with which it classifies a true non-case as a non-case. On the other hand, differential measurement error can cause a bias in any direction, and may even reverse the direction of the association (e.g. RR below 1 instead of above 1).

A good example of the effect of outcome measurement error deals with the risk of soft-tissue sarcomas (STS) among workers exposed to dioxin (see also Chapter 8). STS are rare malignant neoplasms originating in tissues of mesodermal origin, such as fat, muscle and connective tissue. However, although most STS originate from organs primarily containing mesodermal tissue, some originate from visceral organs, such as the stomach, that also contain mesodermal tissue, but such sarcomas represent only a small fraction of all neoplasms originating in these organs. Since the International Classification of Diseases (ICD), which is used in many countries to code causes of deaths from death certificates, is based on the topography of the neoplasms, STS arising from visceral organs are coded as neoplasms of those organs, and are therefore not identified in a mortality study, which will be based on the ICD category for malignant neoplasms of connective and other soft tissues. The net result is a 30–70% underestimation of the number of STS (Lynge *et al.*, 1987; Erikson & Gezelius, 1990).

Exposure to dioxin has been related to an increased risk of STS in a number of cohort studies in which the mortality of the exposed workers has been compared with the mortality of the national population (see section 9.2). In particular, a study of over 6000 workers in the USA has found a threefold excess mortality from STS, based on four observed and 1.2 expected cases, although the result was not statistically significant (Fingerhut *et al.*, 1991). A careful examination of all death certificates of deceased cohort members and of the medical records of all those dyi sites where STS are likely to oc additional cases of STS (Sur Therefore, three[1] out of seven S missed in the analysis based o The degree of misclassificatio population used in this study one can reasonably assume tha tion in the original comparisor tial. The RR of 3.3 (4/1.2), alth then less precise than it woul been possible to use all the cas in the cohort in its estimatio interval based on four cases, seven cases, 1.4–6.9). A narrow val would allow a more precis mation of risk based on these r parison of the seven cases been ulation expected value of 1.2, bly have been differential erro surement, with the RR biased a

4.7 Conclusions

Errors in measurement of come can have substantial and unpredictable effects on the di of dose–response relationship error can be prevented by car design of measurement instru quality control in their use an and analysis of the measurem prevented, it can be controlled the use of multiple measures of come, the adjustment of dose ships for the effects of measure control of covariates related to While the application of thes sometimes make the effects of negligible in QEP, this is probab

There are a number of area could lead to better measure and outcome for QEP. First, dev rate biological techniques to m effective dose may produce es exposure (e.g. DNA adducts) exposure (e.g. exposure-specifi are more accurate than any o of exposure currently available

[1] Two were not STS.

PM3006483550

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

medical records of all those dying from cancers in sites where STS are likely to occur identified three additional cases of STS (Suruda *et al.*, 1993). Therefore, three[1] out of seven STS cases (43%) were missed in the analysis based on death certificates. The degree of misclassification in the reference population used in this study is not known, but one can reasonably assume that the misclassification in the original comparison was non-differential. The RR of 3.3 (4/1.2), although not biased, is then less precise than it would have been had it been possible to use all the cases of STS identified in the cohort in its estimation (95% confidence interval based on four cases, 0.9–8.7; based on seven cases, 1.4–6.9). A narrower confidence interval would allow a more precise quantitative estimation of risk based on these results. Had a comparison of the seven cases been made with the population expected value of 1.2, there would probably have been differential error in outcome measurement, with the RR biased away from the null.

4.7 Conclusions

Errors in measurement of exposure and outcome can have substantial and, to a large degree, unpredictable effects on the direction and shape of dose–response relationships. Measurement error can be prevented by careful selection and design of measurement instruments, and strict quality control in their use and in the processing and analysis of the measurement results. If not prevented, it can be controlled to some degree by the use of multiple measures of exposure and outcome, the adjustment of dose–response relationships for the effects of measurement error, and the control of covariates related to measurement error. While the application of these procedures may sometimes make the effects of measurement error negligible in QEP, this is probably rarely the case.

There are a number of areas of research that could lead to better measurements of exposure and outcome for QEP. First, development of accurate biological techniques to measure biologically effective dose may produce estimates of present exposure (e.g. DNA adducts) or of cumulative exposure (e.g. exposure-specific mutations) that are more accurate than any other measurements of exposure currently available. The use of the first, however, is limited by the short exposure period that they cover, and the use of the second by the fact that they can give essentially no information about dose rate or pattern of exposure. To evaluate these problems and their significance, it would be desirable to incorporate in large cohort studies the periodic collection of biological samples for recurrent measurement of present exposure (to obtain information on dose rate, pattern of exposure and, ultimately, cumulative exposure over some period of time) and the collection, on a case–control basis within a cohort study, of biological measurements of cumulative exposure, ideally corresponding in some way to the measurements of present exposure.

Second, the development, as measures of outcome, of markers of early effects which are important in the process of cancer development may lead to studies that can be carried out sooner after exposure, more quickly, and with greater statistical power than those based on cancer incidence or mortality. The best way to investigate the significance and role of markers of early effect would be to make repeated collections of biological samples in prospective studies, and to carry out large surveys of these markers in high-risk populations with well characterized exposures.

Finally, there would be value in a thorough theoretical and empirical evaluation of the effects of measurement error on dose–response relationships and QEPs derived from them. The theory has been worked out only recently and, as yet, incompletely. Further substantial development and the exploration of its application to some of the more comprehensive data on dose–response in cancer epidemiology are possible.

References

Armstrong, B. (1985) The use of epidemiological data to assess human cancer risk. In: Vouk, V.B., Butler, G.C. & Peakall, D.B., eds, *Methods for Estimating Risk of Chemical Injury: Humans and Non-human Biota and Ecosystems*, Chichester, Wiley, pp. 289-301

Armstrong, B.K. (1988) Epidemiology of malignant melanoma: intermittent or total accumulated exposure to the sun. *J. Dermatol. Surg. Oncol.*, **14**, 835-849

[1] Two were not STS.

PM3006483551

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Armstrong, B.K. & English, D.R. (1988) The epidemiology of acquired melanocytic naevi and their relationship to malignant melanoma. *Pigment Cell*, 9, 27-47

Armstrong, B.K., White, E. & Saracci, R. (1992) *Principles of Exposure Measurement in Epidemiology*, Oxford, Oxford University Press

Birkett, N.J. (1992) Effect of nondifferential misclassification on estimates of odds ratios with multiple levels of exposure. *Am. J. Epidemiol.*, 136, 356-362

Brenner, H. & Blettner, M. (1993) Misclassification bias arising from random error in exposure measurement: Implications for dual measurement strategies. *Am. J. Epidemiol.*, 138, 453-461

Brenner, H., Savitz, D.A., Jöckel, K.-H. & Greenland, S. (1992) Effects of nondifferential exposure misclassification in ecologic studies. *Am. J. Epidemiol.*, 135, 85-95

Brenner, H., Savitz, D.A. & Gefeller, O. (1993) The effects of joint misclassification of exposure and disease on epidemiologic measures of association. *J. Clin. Epidemiol.*, 46, 1195-1202

Brinton, L.A. (1992) Epidemiology of cervical cancer—overview. In: Muñoz, N., Bosch, F.X., Shah, K.V. & Meheus, A., eds, *The Epidemiology of Human Papillomavirus and Cervical Cancer* (IARC Scientific Publications No. 119), Lyon, International Agency for Research on Cancer, pp. 3-23

Chen, J., Campbell, T.C., Li, J. & Peto, R. (1990) *Diet, Lifestyle and Mortality in China: A Study of the Characteristics of 65 Chinese Counties*, Oxford, Oxford University Press

Coghlin, J., Hammond, S.K. & Gann, P.H. (1989) Development of epidemiologic tools for measuring environmental tobacco smoke exposure. *Am. J. Epidemiol.*, 130, 696-704

Copeland, B. (1989) Quality control. In: Kaplan, L.A. & Pesce, A.J., eds, *Clinical Chemistry: Theory, Analysis and Correlation*, 2nd ed., St Louis, MO, Mosby, pp. 279-289

Cummings, K.M., Markello, S.J., Mahoney, M.C & Marshall, J.R. (1989) Measurement of lifetime exposure to passive smoke. *Am. J. Epidemiol.*, 130, 122-132

de Waard, F. (1986) Body size, body mass and cancer of the breast. In: Ip, C., Birt, D.F. & Rogers, A.E., eds, *Dietary Fat and Cancer*, New York, NY, Alan R. Liss, pp. 33-41

Doll, R., Peto, R., Wheatley, K., Gray, R. & Sutherland, I. (1994) Mortality in relation to smoking: 40 years' observations on male British doctors. *Br. Med. J.*, 309, 901-911

Doll, R. & Peto, J. (1978) Cigarette smoking and bronchial carcinoma. Dose and time relationships among regular smokers and lifelong non-smokers. *J. Epidemiol. Commun. Health*, 32, 303-313

Dunn, G. (1989) *Design and analysis of reliability studies*. London, Edward Arnold, and New York, NY, Oxford University Press

English, D.R. & Armstrong, B.K. (1994) The epidemiology of benign melanocytic nevi in children. Observer variation in counting nevi. *Am. J. Epidemiol.*, 139, 402-407

Erikson, A. & Gezelius, C. (1990) Malignant diseases in cause-of-death autopsies in forensic medicine. *Nord. Med.*, 105(10), 264-265

Feinstein, A.R. (1985) A bibliography of publications on observer variability. *J. Chron. Dis.*, 38, 619-632.

Fingerhut, M.A., Halperin, W.E., Marlow, D.A., Piacitelli, L.A., Honchar, P.A., Sweeney, M.H., Greife, A.L., Dill, P.A., Steenland, K. & Suruda, A.J. (1991) Cancer mortality in workers exposed to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin, *New. Engl. J. Med.*, 1991, 324(4): 212-218

Fleiss, J.L. (1971) Measuring nominal scale agreement among many raters. *Psychol. Bull.*, 76, 378-382

Fontham, E.T.H., Correa, P., Wu-Williams, A., Reynolds, P., Greenberg, R.S., Buffler, P.A., Chen,

V.W., Boyd, P., Alterman
Greenberg, S.D. (1991) I
ing women: a multice
Cancer Epidemiol. Bioma

Freedman, L.S., Schatzk
The impact of dietary
planning sample size r
study. *Am. J. Epidemiol.*,

Garfinkel, L. (1981) Tim
mortality among nonsm
sive smoking. *J. Natl Ca*

Gilbert, E.S., Omohund
Holter, N.A. (1993) Mo
Hanford site: 1945–1986.

Greenland, S. (1988) St
to misclassification: im
studies. *J. Clin. Epidemio*

Greenland, S. (1992) Div
ic and individual-level s
1209-1223

Greenland, S. & Mo
Ecological bias, confoun
cation. *Int. J. Epidemiol.*,

Greenland, S. & R
Confounding and mis
Epidemiol., 122, 495-506

Hattis, D. & Silver, K. (199
risk assessment. In: Schu
eds, *Molecular Epidemi*
Practices, San Diego, CA
251-273

Heenan, P.J., Matz, L.R.,
G.R.H., Singh, A., ten Sel
C.D.J. (1984) Inter-obse
pathologists in the class
malignant melanoma i
Histopathology, 6, 581-589

Hiatt, R.A., Krieger, N.,
W.H., Jr & Mihm, M.C. (
and the excess risk of

PM3006483552

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

V.W., Boyd, P., Alterman, T., Austin, D.F., Liff J & Greenberg, S.D. (1991) Lung cancer in nonsmoking women: a multicenter case–control study. *Cancer Epidemiol. Biomarkers Prev.*, 1, 35-43

Freedman, L.S., Schatzkin, A. & Wax, Y. (1990) The impact of dietary measurement error on planning sample size requirements in a cohort study. *Am. J. Epidemiol.*, 132, 1185-1195

Garfinkel, L. (1981) Time trends in lung cancer mortality among nonsmokers and a note on passive smoking. *J. Natl Cancer Inst.*, 66, 1061-1066

Gilbert, E.S., Omohundro, E., Buchanan, J. & Holter, N.A. (1993) Mortality of workers at the Hanford site: 1945–1986. *Health Phys.*, 64, 577-590

Greenland, S. (1988) Statistical uncertainty due to misclassification: importance for validation studies. *J. Clin. Epidemiol.*, 41, 1167-1174

Greenland, S. (1992) Divergent biases in ecologic and individual-level studies. *Stat. Med.*, 11, 1209-1223

Greenland, S. & Morgenstern, H. (1989) Ecological bias, confounding, and effect modification. *Int. J. Epidemiol.*, 18, 269-274

Greenland, S. & Robins, J.M. (1985) Confounding and misclassification. *Am. J. Epidemiol.*, 122, 495-506

Hattis, D. & Silver, K. (1993) Use of biomarkers in risk assessment. In: Schulte, P.A. & Perera, F.P., eds, *Molecular Epidemiology. Principles and Practices*, San Diego, CA, Academic Press, pp. 251-273

Heenan, P.J., Matz, L.R., Blackwell, J.B., Kelsall, G.R.H., Singh, A., ten Seldam, R.E.J. & Holman, C.D.J. (1984) Inter-observer variation between pathologists in the classification of cutaneous malignant melanoma in Western Australia. *Histopathology*, 6, 581-589

Hiatt, R.A., Krieger, N., Sagebiel, R.W., Clark, W.H., Jr & Mihm, M.C. (1993) Surveillance bias and the excess risk of malignant melanoma among employees of the Lawrence Livermore National Laboratory. *Epidemiology*, 4, 43-47

Higgins, I.T.T, Mahan, C.M. & Wynder, E.L. (1988) Lung cancer among cigar and pipe smokers. *Prev. Med.*, 17, 116-128

Hirayama, T. (1981) Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study from Japan. *Br. Med. J.*, 282, 183-185

Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C.C. (1991) *p53* mutations in human cancers. *Science*, 253(5015): 49-53

Holman, C.D.J., Armstrong, B.K. & Heenan, P.J. (1983) A theory of the etiology and pathogenesis of human cutaneous malignant melanoma. *J. Natl Cancer Inst.*, 71, 651-656

Hulka, B.S., Wilcosky, T.C. & Griffith, J.D. (1990) *Biological Markers in Epidemiology*, New York, NY, Oxford University Press

Jones, P.A., Buckley, J.D., Henderson, B.E., Ross, R.K. & Pike, M.C. (1991) From gene to carcinogen: a rapidly evolving field in molecular epidemiology. *Cancer Res.*, 51(13), 3617-3620

Kaldor, J.M., Day, N.E., Clarke, E.A., van Leeuwen, F.E., Henry-Amar, M., Fiorentino, M.V., Bell, J., Pedersen, D., Band, P., Assouline, D. *et al.* (1990) Leukemia following Hodgkin's disease. *New Engl. J. Med.*, 322, 7-13

Kraemer, H.C. (1992) How many raters? Toward the most reliable diagnostic consensus. *Stat. Med.*, 11, 317-331

Kristensen, P. (1992) Bias from non-differential but dependent misclassification of exposure and outcome. *Epidemiology*, 3, 210-215

Landis, J.R. & Koch, G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174

Lee-Feldstein, A. (1989) A comparison of several measures of exposure to arsenic. *Am. J. Epidemiol.*, 129, 112-124

PM3006483553

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Liu, K., Cooper, R., McKeever, J., McKeever, P., Byington, R., Soltero, I., Stamler, R., Gosch, F., Stevens, E. & Stamler, J. (1979) Assessment of the association between habitual salt intake and high blood pressure: methodological problems. *Am. J. Epidemiol.*, **110**, 219-226

Lynge, E., Storm, H.H. & Jensen, O.M. (1987) The evaluation of trends in soft tissue sarcoma according to diagnostic criteria and consumption of phenoxy herbicides. *Cancer*, **60**, 1896-1901

Maclure, M. & Willett, W.C. (1987) Misinterpretation and misuse of the kappa statistic. *Am. J. Epidemiol.*, **126**, 161-169

MacMahon, B. & Pugh, T.F. (1970) *Epidemiology. Principles and Methods*, Boston, MA, Little, Brown

Marshall, J.R. (1989) The use of dual or multiple reports in epidemiologic studies. *Stat. Med.*, **8**, 1041-1049

Mazoyer, B. & Mary, J.T. (1987) Kappa as an index of reproducibility: Distribution under the null-hypothesis. *Rév. Epidemiol. Santé Publ.*, **35**, 474-481

Muñoz, N., Bosch, F.X., de Sanjosé, S., Tafur, L., Izarzugaza, I., Gili, M., Viladiu, P., Navarro, C., Martos, C., Ascunce, N., Gonzalez, L.C., Kaldor, J.M., Guerrero, E., Lorincz, A., Santamaria, M., Alonso de Ruiz, P., Aristizabal, N. & Shah, K. (1992) The causal link between human papillomavirus and invasive cervical cancer: a population-based case–control study in Colombia and Spain. *Int. J. Cancer*, **52**, 743-749

Nakazawa, H., English, D., Randell, P.L., Nakazawa, K., Martel, N., Armstrong, B.K. & Yamasaki, H. (1994) UV and skin cancer: specific *p53* gene mutation in normal skin as a biologically relevant exposure measurement. *Proc. Natl Acad. Sci. USA*, **91**, 360-364

National Institutes of Health (1985) *Report of the National Institutes of Health Ad Hoc Working Group to Develop Radioepidemiological Tables* (NIH Publication No. 95-2748), Washington, DC, US Department of Health and Human Services

Parkin, D.M. & Muir, C. (1992) Comparability and quality of data. In: *Cancer Incidence in Five Continents*, Vol. VI (IARC Scientific Publications No. 120), Lyon, International Agency for Research on Cancer, pp 45-173

Parkin, D.M., Muir, C.S., Whelan, S.L., Gao, Y-T., Ferlay, J. & Powell, J., eds (1992) *Cancer Incidence in Five Continents*, Vol. VI (IARC Scientific Publications No. 120), Lyon, International Agency for Research on Cancer

Percy, C. & Muir, C. (1989) The international comparability of cancer mortality data. *Am. J. Epidemiol.*, **129**, 934-946

Perera, F.P. & Santella, R. (1993) Carcinogenesis. In: Schulte, P.A. & Perera, F.P., eds, *Molecular Epidemiology. Principles and Practices*, San Diego, CA, Academic Press, pp. 277-300

Perera, F.P. & Weinstein, I.B. (1982) Molecular epidemiology and carcinogen–DNA adduct detection—new approaches to studies of human cancer causation. *J. Chron. Dis.*, **35**, 581-600

Perera, F., Mayer, J., Santella, R.M., Brenner, D., Jeffrey, A., Latriano, L., Smith, S., Warburton, D., Young, T.L., Tsai, W.Y., Hemminki, K. & Brandt-Rauf, P. (1991) Biologic markers in risk assessment for environmental carcinogens. *Environ. Health. Perspect.*, **90**, 247-254

Persson, P.-G. & Norell, S.E. (1989) Retrospective versus original information on cigarette smoking. Implications for epidemiological studies. *Am. J. Epidemiol.*, **130**, 705-712

Phillips, A.M. & Davey Smith, G. (1993) The design of prospective epidemiological studies: more subjects or better measurement. *J. Clin. Epidemiol.*, **46**, 1203-1211

Pierce, D.A., Stram, D.O. & Vaeth, M. (1990) Allowing for random errors in radiation exposure estimates for the atomic bomb survivor data. *Radiat. Res.*, **123**: 275-284

Pierce, D.A., Stram, D.O., Vaeth, M. & Schafer, D.W. (1992) The errors in variables problem: con-

siderations provided by analyses of the A-bomb *Assoc.*, 87:351-359

Preston, D.L. & Pierce, of changes in dosimet risk estimates in the a *Radiat. Res.*, **114**, 437-4

Preston, D.L., Kato, H., (1987) Studies of the m vivors. 8. Cancer morta *Res.*, **111**, 151-178

Riboli, E., Preston-Marti N.J., Trichopoulos, D., Fontham, E.T.H., Gao, L.C., Le Marchand, L., S Stanta, G., Wu-William (1990) Exposure of nons ronmental tobacco smok rative study. *Cancer Cau*

Rogot, E. & Reid, D.D. data from next-of-kin i among migrants. *Int. J.*

Rosner B., Willett, W. (1989) Correction of logi risk estimates and confi tematic within-person *Stat. Med.*, **8**, 1051-1069

Rothman, K.J. (1981) periods. *Am. J. Epidemiol*

Schiffman, M.H. Bauer, Manos, M.M., Byrne, J. D.M. & Howley, P.M. (Southern blot hybridiza chain reaction methods human papillomaviru *Microbiol.*, **29**, 573-577

Schulte, P.A., Rothman, (1993) Design conside epidemiology. In: Schult eds, *Molecular Epidemi Practices*, San Diego, CA, 159-198

PM3006483554

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

siderations provided by radiation dose–response analyses of the A-bomb survivor data. *J. Am. Stat. Assoc.*, 87:351-359

Preston, D.L. & Pierce, D.A. (1988) The effect of changes in dosimetry on cancer mortality risk estimates in the atomic bomb survivors. *Radiat. Res.*, **114**, 437-466

Preston, D.L., Kato, H., Kopecky, K. & Fujita, S. (1987) Studies of the mortality of A-bomb survivors. 8. Cancer mortality, 1950–1982. *Radiat. Res.*, **111**, 151-178

Riboli, E., Preston-Martin, S., Saracci, R., Haley, N.J., Trichopoulos, D., Becher, H., Burch, J.D., Fontham, E.T.H., Gao, Y.-T., Jindal, S.K., Koo, L.C., Le Marchand, L., Segnan, N., Shimizu, H., Stanta, G., Wu-Williams, A.H. & Zatonski, W. (1990) Exposure of nonsmoking women to environmental tobacco smoke: a 10-country collaborative study. *Cancer Causes Control*, 1, 243-252.

Rogot, E. & Reid, D.D. (1975) The validity of data from next-of-kin in studies of mortality among migrants. *Int. J. Epidemiol.*, 4, 51-54

Rosner B., Willett, W.C. & Spiegelman, D. (1989) Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat. Med.*, 8, 1051-1069

Rothman, K.J. (1981) Induction and latent periods. *Am. J. Epidemiol.*, 114, 253-259

Schiffman, M.H. Bauer, H.M. Lörincz, A.T., Manos, M.M., Byrne, J.C., Glass, A.G., Cadell, D.M. & Howley, P.M. (1991) Comparison of Southern blot hybridization and polymerase chain reaction methods for the detection of human papillomavirus DNA. *J. Clin. Microbiol.*, 29, 573-577

Schulte, P.A., Rothman, N., & Schottenfeld, D. (1993) Design considerations in molecular epidemiology. In: Schulte, P.A. & Perera, F.P., eds, *Molecular Epidemiology. Principles and Practices*, San Diego, CA, Academic Press, pp. 159-198

Schwartz, G.C. (1990) Chromosome aberrations. In: Hulka, B.S., Wilcosky, T.C. & Griffith, J.D., eds, *Biological Markers in Epidemiology*, New York, NY, Oxford University Press, pp. 147-172

Shimizu, Y., Schull, W.J. & Kato, H. (1990) Cancer risk among atomic bomb survivors. The RERF Life Span Study. Radiation Effects Research Foundation comments. *J. Am. Med. Soc.*, **264**, 601-604

Smith, A.H., Waxweiler, R.J. & Tyroler, H.A. (1980) Epidemiologic investigation of occupational carcinogenesis using a serially additive expected dose model. *Am. J. Epidemiol.*, **112**, 787-797

Snedecor, G.W. & Cochran, W.G. (1980) *Statistical Methods*, Ames, IA, Iowa State University Press, pp. 243-244

Spiegelman, D. & Gray, R. (1991) Cost efficient designs for binary responses data with Gaussian covariate measurement error. *Biometrics*, **47**, 851-869

Steenland, K. & Schnorr, T. (1988) Availability and accuracy of cancer and smoking data obtained from next of kin for decedents in a retrospective cohort study. *J. Occup. Med.*, **30**, 348-353

Stockwell, H.G., Goldman, A.L., Lyman, G.H., Noss, C.I., Armstrong, A.W., Pinkham, P.A., Candelora, E.C. & Brusa, M.R. (1992) Environmental tobacco smoke and lung cancer risk in nonsmoking women. *J. Natl Cancer Inst.*, **84**, 1417-1422

Straume, T., Egbert, S.D., Woolson, W.A., Finkel, R.C., Kubik, P.W., Gove, H.E., Sharma, P. & Hoshi, M. (1992) Neutron discrepancies in the DS86 Hiroshima dosimetry system. *Health Phys.*, 63:421-426

Suruda, A.J., Ward, E.M. & Fingerhut, M.A. (1993) Identification of soft tissue sarcoma deaths in cohorts exposed to dioxin and to chlorinated naphthalenes. *Epidemiology*, **4**, 14-19

PM3006483555

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001

Thomas, D.C., Strom, D. & Dwyer, J. (1993) Exposure measurement error: Influence on exposure–disease relationships and methods of correction. *Ann. Rev. Public Health*, 14, 69-93

UK National Case–Control Study Group (1989) Oral contraceptive use and breast cancer risk in young women. *Lancet*, i, 973-982

Vine, M.F. (1990) Micronuclei. In: Hulka, B.S., Wilcosky, T.C. & Griffith, J.D., eds, *Biological Markers in Epidemiology*, New York, NY, Oxford University Press, pp. 125-146

Vineis, P., Schulte, P.A. & Vogt, R.F. (1993) Technical variability in laboratory data. In: Schulte, P.A. & Perera, F.P., eds, *Molecular Epidemiology. Principles and Practices*, San Diego, CA, Academic Press, pp. 109-135

Wacholder, S., Armstrong, B. & Hartge, P. (1993) Validation studies using a gold standard. *Am. J. Epidemiol.*, 137, 1251-1258

Walker, A.M., Velema, J.P. & Robins, J.M. (1988) Analysis of case–control data derived in part from proxy respondents. *Am. J. Epidemiol.*, 127, 905-914

Westgard, J.O. & Klee, G.G. (1986) Quality assurance. In: Tietz, N.W., ed., *Textbook of Clinical Chemistry*, Philadelphia, PA, Saunders, pp. 424-458

Wilcosky, T.C. & Rynard, S.M. (1990) Sister chromatid exchange. In: Hulka, B.S., Wilcosky, T.C. & Griffith, J.D., eds, *Biological Markers in Epidemiology*, New York, NY, Oxford University Press, pp. 105-124

Wild, C.P., Jiang Y.-Z., Sabbioni, G., Chapot, B. & Montesano, R. (1990) Evaluation of methods for quantitation of aflatoxin–albumin adducts and their application to human exposure assessment. *Cancer Res.*, 50, 245-251

Willett, W.C. (1989) An overview of issues related to the correction of non-differential exposure measurement error in epidemiologic studies. *Stat. Med.*, 8, 1031-1040

Ziegler, A., Leffell, D.J., Kunala, S., Sharma, H.W., Gailani, M., Simon, J.A., Halpern, A.J., Baden, H.P. Shapiro, P.E., Bale, A.E. & Brash, D.E. (1993) Mutation hotspots due to sunlight in the *p53* gene of nonmelanoma skin cancers. *Proc. Natl Acad. Sci. USA*, 90, 4216-4220

5: Long- a Carcinoge and Short-

Victor J. Feron, Michael

5.1 General principles

The primary objective c bioassays has been to identif induce cancer in humans. I human data, studies in la (rodents) still remain the mos widely accepted means for th cinogenic hazards to public h 1992; Griesemer, 1992; Vaini addition to the screening of c cinogenic potential on a qual cinogenicity study may be e mine dose–response relationsl the mechanism of carcinoge *al.*, 1986; Krewski *et al.*, 1992) studies are performed mainly cinogenic potential of the te determine dose–response relat agent studies are focused o mechanisms of action, viz. i ing, inhibiting or cocarcinoge

Studies that examine certai cerning mechanisms of ca designs tailored to a partic classical example of specific tion–promotion studies that u mouse skin or rat liver as exp Other examples are two-gen distinguish between effects i or postnatally, studies with c exposure periods of varying d (ir)reversibility of hyper/pren and mixed-exposure studies t sibility of synergistic or anta (Gart *et al.*, 1986; Griesemer Vainio *et al.*, 1992). Importan as dose– and time–response f in target and non-target tiss *al.*, 1991), and presence or ab

PM3006483556

Source: https://www.industrydocuments.ucsf.edu/docs/txxj0001